DEVELOPMENT OF A NEW APPROACH FOR PREDICTING TRAM TRACK DEGRADATION BASED ON PASSENGER RIDE COMFORT DATA

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Amir Falamarzi

B.Eng. (Civil Eng), Azad University, Karaj, Iran

M.Eng. (Transport Engineering), Azad University, Tehran, Iran

School of Engineering

College of Science, Engineering and Health

RMIT University

January 2020

Declaration

I certify that except where the acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work; paid or unpaid; carried out by a third party is acknowledged; and; ethics procedures and guidelines have been followed. Furthermore, I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Amir Falamarzi

7 January 2020

Acknowledgements

I would like to express my sincere and deep gratitude to my main supervisor, Sara Moridpour, who has kindly provided me with an opportunity to do PhD under her supervision. She helped me with her feedback and technical support in different stage of this work. She taught me the methods and tools to carry out this research and prepare it as accurate and practical as possible. It was a great honour to study under her guidance.

I am also grateful to Associate Professor Majidreza Nazem my associate supervisor who his advice and recommendations were effective in finalising my PhD. I learnt so many useful things from his feedback and positive criticisms. He gave me constructive ideas during my studies and has always encouraged me to explore new ideas.

My sincere thanks go to all the academic and administrative staff in School of Engineering at RMIT University. I would like to thank RMIT University for supporting my APA scholarship.

I would like to thank my wife, Samira Cheraghi, for her support and help. She is the person behind my success and I am indebted to her sacrifice and understanding. Words would never say how grateful I am to my wife.

Finally, I would like to pay my regards and respects to my beloved parents for their unconditional support and love.

List of Publications and Awards

Journal Papers:

1. Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2018). Rail degradation prediction models for tram system: Melbourne case study. *Journal of Advanced Transportation*, 2018. (Q1)
2. Falamarzi, A., Moridpour, S., & Nazem, M. (2019). Development of a tram track degradation prediction model based on the acceleration data. *Structure and Infrastructure Engineering*, 1-11. (Q1)
3. Falamarzi, A., Moridpour, S., & Nazem, M. (2019). A time-based track quality index: Melbourne tram case study. *International Journal of Rail Transportation*. (Q1)
4. Falamarzi, A., Moridpour, S., & Nazem, M. (2019). A review of rail track degradation prediction models. *Australian Journal of Civil Engineering*, 17(2), 152-166. (Q3)
5. Falamarzi, A., Moridpour, S., Nazem, M., & Cheraghi, S. (2019). Prediction of tram track gauge deviation using artificial neural network and support vector regression. *Australian Journal of Civil Engineering*, 1-9. (Q3)
6. Falamarzi, A., Moridpour, S., & Nazem, M. (2019). A Review on Existing Sensors and Devices for Inspecting Railway Infrastructure. *Jurnal Kejuruteraan*, 31(1), 1-10.

Peer-review Conference Papers:

1. Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2017). Rail degradation predication: Melbourne tram system case study. *Australian Transport Research Forum (ATRF) 39th*, Auckland, New Zealand.
2. Falamarzi, A., Moridpour, S., Nazem, M., & Cheraghi, S. (2018). Development of Random forests regression model to predict track degradation index: Melbourne case study. *Australasian Transport Research Forum (ATRF) 40th*, Darwin, Northern Territory, Australia.

3. Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2019). Integration of genetic algorithm and support vector machine to predict rail track degradation. *MATEC Web of Conferences* (Vol. 259, p. 02007). EDP Sciences.

Awards:

1. RMIT Engineering HDR Publication Award 2018 for "Rail degradation prediction models for tram system: Melbourne case study. *Journal of Advanced Transportation*, 2018".
2. RMIT Engineering HDR Publication Award 2019 for "Development of a tram track degradation prediction model based on the acceleration data. *Structure and Infrastructure Engineering*, 1-11".
3. RMIT Post-submission Publication Award 2020.

Table of Contents

List of Figures

List of Tables

Summary

These days tram as a type of the public transport system has become popular because of its attractive features such as road usage efficiency, low emission of pollutants, reduction in traffic congestion and efficiency in capital costs and maintenance expenses compared to private cars. For the case study, the Melbourne tram network, which is the longest tram network in the world, has been targeted. Melbourne tram system consists of 493 trams, 24 routes, and 1,763 tram stops. According to the operator of the Melbourne tram network, the total number of patronage in 2017-2018 was 206.3 million. In parallel with the annual increase in tram demand and patronage, tram infrastructure systems need to bear more stresses and traffic pressure. Track degradation is a common problem in the area of tram track infrastructure. One of the main aspects of track degradation is the presence of irregularity in track geometric parameters.

In order to deal with degradation problems, tram track infrastructure maintenance management systems have been developed for design and implementation of maintenance works and renewal activities. Track degradation prediction models are the core and the main part of these management systems. Without accurately predicting the future condition of tram tracks, designing and providing preventive maintenance strategies are not feasible. In this research, the collected data which cover six sequential years (2010 to 2015) have been analysed and influencing parameters in tram track degradation have been identified. Gauge and twist were identified as the influencing track geometry parameters in the tram track degradation. Besides that, track surface and rail support as structural parameters were identified as significant parameters in prediction of future track geometry parameters and consequently tram track degradation.

In order to develop tram track degradation prediction models and according to the successful experience of the previous studies, three types of prediction models including Artificial Neural Network (ANN), Support Vector Machine (SVM) and Random Forest Regression (RFR) models have been created. According to the

results, RFR models provide better predictions in terms of the performance indicators including the coefficient of determination and Root Mean Squared Error (RMSE) compared to the ANN and SVM models.

In this research, based on the Melbourne tram track dataset, a new track degradation index has been proposed. Track degradation indices can be used as an indicator of rail condition concerning the risk of damage or failure over a period of time. The index can be applied in establishing a sustainable tram track maintenance management system. The new index composed of two main parts including the mean value of the geometry deviation and the average differential geometry deviation. The proposed index has been compared with three major track geometry degradation indices. For this purpose, the predictability performance of the indices has been considered. In this regard, the Pearson correlation analysis was applied to previous and current values of the indices. According to the results, the correlation coefficient of the proposed index was higher than the other indices. The finding of the evaluation presented that the proposed index can be used as an effective measure for the assessment of the geometric condition of tram tracks.

In this research, a new approach has been proposed to predict the tram track degradation were which is cost-effective and can be carried out repeatedly without imposing delay to tram services. Conventional approaches are mainly based on the previous track geometry parameters which have been discussed in this research. In the new approach, passenger ride comfort data or acceleration data has been used to predict the future condition of track geometry parameters which has been represented by the tram track degradation index. For developing the degradation prediction models, the previous models which have been used to predict the degradation based on the track geometry parameters were applied. The future degradation index has been targeted as the target variable and acceleration parameter besides the structural parameters have been used as the explanatory variables. According to the results of the evaluation, the RFR model can predict the future degradation index with approximately 10 percent higher R^2 and 9 percent lower prediction error compared to other developed models.

In this research two methods for predicting the future tram track degradation index, first was the method based on the previous track geometry parameters and the second was the method based on the acceleration data, have been presented. According to the results of the degradation index prediction based on the previous track geometry parameters, RMSE was 0.35 and R^2 value was 0.95. On the other hand, for the prediction based on the acceleration data, RMSE was 1.04 and R^2 value was 0.74. The comparison of these methods shows that although the prediction error has been increased and R^2 value has been decreased in the latest method, the values of the performance indicators are still in acceptable ranges. These results imply that the prediction of tram track degradation based on the acceleration data can be considered as a reliable method along with conventional tram track degradation prediction method for maintaining tram tracks. The proposed method can provide more predictions of potential future faults by reducing inspection costs and inspection intervals.

CHAPTER 1

INTRODUCTION

1.1. Background

Railway transport can be divided into two main categories including heavy rail land light rail. Rapid transit and suburban rail system are considered as the main subcategories of heavy rail. Light Rail Transit (LRT) and tram are mainly categorised under light rail systems. Trams as urban rail transport are used in a different part of the world to accommodate the movement of people within suburbs and cities. Tram system have evolved gradually during the years after its introduction. Trams can be divided into several types in terms of the power used including horse-drawn trams, steam trams, cable-hauled trams, gas trams and electric trams. Nowadays, electric trams are the most common type of trams that have been using in different places (De Bruijn & Veeneman 2009).

Tram as a means of the public transport system has some general advantages compared to private vehicles. In terms of road usage efficiency, trams can take up far less spaces of the road compared to cars while carrying more passengers. In terms of reduction in air pollutions using tram systems is more efficient than private vehicles. In terms of reduction in traffic congestion, tram systems can ease traffic-related problems in urban areas. Different examples of implementation of tram systems have demonstrated that traffic safety in urban areas has been improved in terms of the number of accidents and casualties. In terms of capital costs and maintenance expenses, trams are more efficient than private cars (Dincer, Hogerwaard & Zamfirescu 2015).

Tram also has some specific advantages among other rail public transport systems. Trams are more accessible for passengers in the heart of cities and can share the road with other vehicles especially when they are compared with underground and elevated rail systems. Due to shape and the structure of tram systems, they are more suitable on tighter curves and higher gradients. Compared to conventional buses, trams can carry more passengers and have higher capacity and in some cities, normal busses are replaced with tram systems. Trams can be used as a public transport system to promote tourism in ways normal buses are not used. (Knowles & Ferbrache 2016; Naznin, Currie & Logan 2016; Diemer et al. 2018).

On the other hand, tram systems have some disadvantages. Tram can be hazardous for cyclist where their tyres can be caught in track grooves. Tram tracks can be dangerous for cyclists and motorcycles when are wet and slippery. Tram systems can also expose the neighbouring populations of the tram tracks to moderate level of noise which can be managed by noise mitigation strategies.

Along with the annual increase in tram demand and patronage, tram infrastructure systems need to bear faster and heavier traffic loads. Having more frequencies of tram service, using new high-speed tram vehicles which are heavier than normal tram vehicles and sharing the route with other vehicles will result in higher rates of tram track degradation. Traffic load has a direct relationship with the existence as well as the length of tracks. Track degradation is a common operational and safety problem in the area of tram track infrastructure (Gaudry, Lapeyre & Quinet 2016; Jamshidi et al. 2017).

One of the main aspects of track degradation is the presence of irregularity in track geometric parameters. Geometry condition of rail track which is measurable should always be kept in an acceptable range. Although the degradation rate of tram track infrastructure evolves gradually, exceeding the acceptable range of geometry parameters can significantly decrease the safety and reliability of tram track infrastructure. Failing to deploy effective preventive track maintenance strategies can lead to disruption in regular tram services as a result of rail track failure or tram derailment (Higgins & Liu 2018).

Tram Track Infrastructure Maintenance Management Systems (TTIMMS) have emerged for design and implementation of maintenance and renewal activities. These systems are necessary to optimise the maintenance of tracks and facilitate effective management of tram infrastructure facilities (Jovanović, Guler & Čoko 2015). Main practices required for a TTRIMMS can be split into different categories. First monitoring and inspection of tram track components. In this context, regular measurement of track geometry parameters is undertaken to compare the exiting deviation with standard and acceptable range. Secondly, based on the data collected from the first step, track degradation prediction models are developed. By applying these models, the future condition of rail tracks in terms of expected geometry irregularities can be revealed. Thirdly, by applying the developed prediction models, track operators can address effective short/long term maintenance and renewal strategies to revitalise tram track infrastructure (Santos, Teixeira & Antunes 2015; Odolinski & Smith 2016).

With regard to the above, prediction modelling of tram track degradation is the fundamental prerequisite for developing efficient and cost-effective maintenance strategies of a tram system. It is evident that without accurately predicting the future condition of tram tracks, designing and providing preventive maintenance strategies are not conceivable (Thaduri, Galar & Kumar 2015)

1.2. Research Questions

Tram track degradation prediction modelling can be considered as a core component of future preventive maintenance activities. There are several questions that need to take into consideration while research is undertaken in the area of tram track degradation prediction modelling. These question then can be applied to develop the research aim and objectives. The main research questions can be outlined as follows:

1. What are the influencing factors in tram track degradation?
2. What type of models can be applied to predict tram track degradation?

3. How different geometry parameters can be integrated into an index to better represent the degradation process in tram track infrastructure?
4. What are the alternative solutions for measuring and monitoring tram track geometry parameters?
5. How can cost-effective approaches be applied to predict the rate of future tram track degradation in tram tracks?

1.3. Research Aim and Objectives

The aim of this research is to develop a cost-effective and efficient method to predict the future condition of tram track geometry parameters. Numerous studies have been carried out in terms of using a cost-effective method to monitor rail track irregularities. Cost-effectiveness can be applied to both method and application that will be used to monitor or predict the future condition of track geometry parameters. In this context, alternative ways to represent track geometry parameters, innovative application and devices that can be used to capture track geometry parameters and efficient models to predict complex dataset are required to be investigated. With regard to the research aim, the main objectives of this research can be listed as follows:

1. To identify the influencing factors contributed to tram track degradation. These factors can be expressed as geometry parameters, structural parameters and other parameters that can illustrate the process of the degradation in tram networks.
2. To predict track geometry degradation based on the existing track geometry and structural parameters. For this purpose, different degradation prediction models are developed. Afterwards, the results of the models are evaluated against each other. The outcome of the evaluation will be used as a benchmark to compare different methods that can be utilised in rail track degradation prediction modelling.
3. To develop an index which can represent the rate of degradation in track geometry parameters over a specific period of time. Furthermore, instead of

using a single track geometry parameter, a track degradation index can be used to represent all the effective parameters together. The proposed index then can be applied by track degradation prediction models for predicting the future condition of tram tracks.

4. To utilise the ride comfort data as a new approach and dataset to monitor tram track geometry degradation. Ride comfort can be measured and represented by acceleration data. This data extracted from the movement of tram on tram tracks will be used as a complement to track geometry measurement.
5. To develop a degradation prediction model based on track degradation index and acceleration data. In the proposed model, vehicle acceleration data along with other track structural parameters are used to predict the degradation index of tram track.
6. To compare two degradation prediction approaches to predict the tram track degradation index. First, the approach based on the modelling of existing track geometry parameters. Second, the approach based on the application of the acceleration data.

1.4. Contribution and Research Scope

The innovative contributions of this research can be listed as follows:

1. Numerous studies have been carried out in relation to heavy rail degradation prediction and modelling, but few studies and experiments have been conducted to model tram track degradation and this study attempts to fill this gap.
2. Current prediction models that are applied for predicting tram track degradation are mostly based on statistical models. One of the contributions of this research is to apply machine learning model on tram track dataset which has not been applied so far. Machine learning models can handle big datasets with complex distribution patterns.
3. Current track degradation indices are developed based on the dataset related to heavy rail infrastructure. In this research, a degradation index will be developed based on the geometry parameters involved in tram tracks.

4. Acceleration data are currently used by rail track maintenance engineers to monitor track irregularities. The contribution of this research is to employ acceleration data for predicting the future condition of rail tracks.

1.5. Benefits of This Research

1. Data collection costs will be lower as acceleration signals can be used as an alternative to tram track geometry data. Few technicians are required for measuring and collecting data as Condition Monitoring Systems (CMS) and nowadays smartphones provide easier ways for these purposes.
2. By using innovative measures such as on-board CMSs and smartphones the process of data collection can be carried out and repeated for several times during a year. Consequently, tram track can be inspected several times instead of limited times.
3. By mounting acceleration sensors on in-service vehicles, the process of vehicle acceleration data collection can be accomplished without any disruption to tram transport services.

1.6. Research Scope

In this research for the case study, the dataset of the Melbourne tram network is used and examined. This dataset which covers both track geometry parameters and the acceleration data is used throughout the research to fulfil the objectives and finally achieve the aim of the research. It must be noted that heavy rail tracks and their associated parameters are not covered in the proposed dataset.

1.7. Thesis Outline

This thesis is structured as follows. Chapter 2 reviews the existing literature in the field of rail track degradation models and also track quality indices which are used in track degradation prediction studies. This chapter examines different types of track degradation models as well as track quality indices with different formulations. Following this research, different parameters that are effective in track degradation can be identified. Based on reviewing the literature, two literature review papers have been published in the *Australian Journal of Civil Engineering and Journal of Engineering*.

Chapter 3 describes the case study, dataset related to the case study and the research framework of this research. This chapter investigates the effective parameters in tram track degradation models and their contribution to the research. Following that, this chapter maps out the research framework of this study. The steps and methods that are required to achieve the aim and objectives of the research.

Chapter 4 describes the track degradation prediction models that are developed based on the dataset of this study which has been represented in Chapter 3. This chapter explores the application of Artificial Intelligence models on the dataset of this research. Furthermore, this chapter reveals the performance of these models on the prediction of tram track degradation. The results of this chapter have been published in the *Journal of Advanced Transportation* and *Australian Transport Research Forum*.

Chapter 5 represents the process of development of a track degradation index based on tram track dataset. The findings of this chapter have been published in the *International Journal of Rail Transportation*. In addition to the index development, this chapter compares the predictability performance of the proposed index and the indices explained in the literature review section.

Chapter 6, first describes the process of the preparation of a new dataset based on the tram track quality index and the acceleration data. Following that, the predication of tram track quality index based on the acceleration data is presented. The results of this chapter have been published in the *Journal of Structure and Infrastructure Engineering*.

Chapter 7 describes the findings of this research. This chapter presents the contribution of this research in terms of the development of the tram track degradation prediction, tram track quality index and the development of degradation prediction models based on the acceleration data. Finally, this chapter ends with future research direction and recommendations toward improving the future datasets and the proposed models.

1.8. Summary

The patronage of tram networks as a public mode of transport is on the rise (PTV 2018). Along with this increase, pressure on the tram infrastructure is also increasing which can lead to tram track degradation. Track geometry degradation is one of the most common causes of rail track failure and tram derailment. To mitigate the risk of the rate of tram track degradation, preventive maintenance measures have been introduced by rail track maintenance management systems. Track degradation prediction modelling is considered as one of the important parts of the preventive maintenance activities. Without applying tram track degradation models, providing accurate and effective maintenance strategies is not feasible. In this chapter, the research questions were outlined. The questions provided in this chapter than were utilised to shape the research aim and accordingly the research objectives.

CHAPTER 2

REVIEW OF THE EXISTING STUDIES

2.1. Introduction

In this section, the literature related to exiting track degradation indices and also track degradation prediction modelling approaches is provided. Track degradation indices as track quality representatives are important as they are mostly created based on data collected over several years. In this section, different track degradation indices based on track geometry parameters used in different researches are discussed. In the second part, track degradation prediction models are investigated.

Degradation prediction modelling is the key element in the establishment of cost-effective and efficient maintenance strategies in railway systems. Modelling the degradation process of rail track geometry parameters is one of the main concerns of railway infrastructure organisers. Within the last decades, various type of track geometry degradation modelling approaches has been developed and applied to predict the future rail track geometry condition. In this section different and important track geometry degradation prediction models are investigated, classified and analysed. Besides that, track geometry degradation indices which can be applied by the above models to represent the future of the track geometry parameters are investigated and summarised. In the end, the summary of the indices and models investigated in this section including their limitations and also parameters or indices involved in the model development are represented.

2.2. Degradation Models

Investigating the railway literature indicates that rail track degradation prediction models can be divided into three main categories: statistical models, mechanistic models and AI models, as illustrated in Figure 2.1. In the following sections, descriptions and examples of the various degradation models are provided

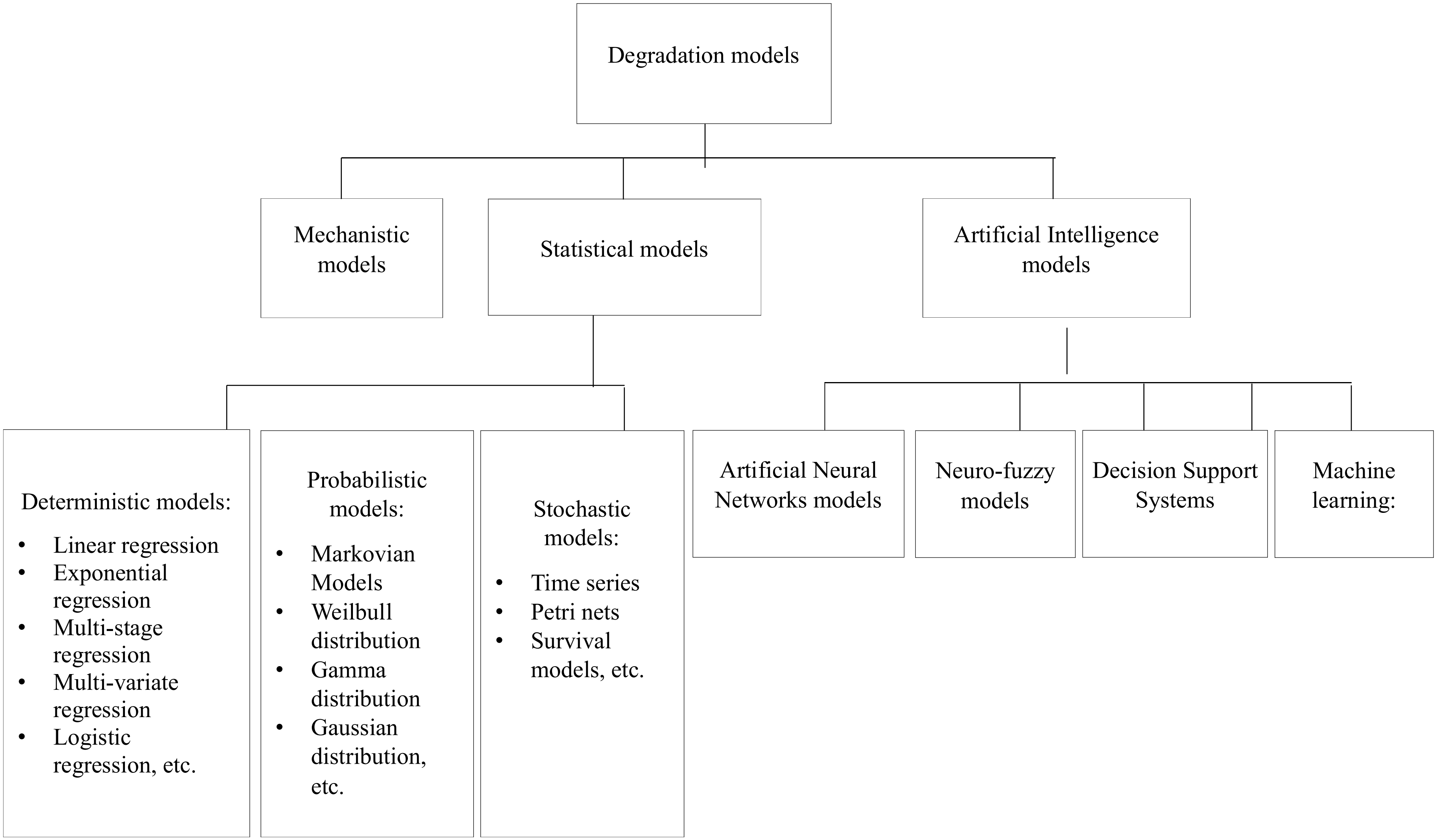

Figure 2.1: Categories of degradation prediction models

2.2.1. Mechanistic Models

Mechanistic models are the oldest models for predicting degradation in railway tracks and can be divided into two categories: conventional mechanistic models and empirical mechanistic models. Conventional mechanistic models can predict track degradation with a small amount of geometrical data. Empirical mechanistic models are a combination of mechanistic and statistical models and because they consider observations and extensive data records, are able to predict the degradation of the entire network rather than a section. Due to their similarity to statistical models, examples of this model are provided in the section for statistical models. In the following, some successful examples of the mechanistic model in track degradation prediction are presented.

The Technical University of Munich has conducted some experiments in the laboratory environment to measure the rate of settlement (s) calculated as follows (Demharter 1982):

$$s = a \times p \times \ln \Delta N + b \times p^{1.21} \times \ln N \qquad (2.1)$$

Where, s is the average rate of irregularity (mm/100 days), p is the ballast pressure (N/mm^2), and ΔN denotes a pre-loading period in addition to the first passing axles. N in the second term is the total number of passing axles. The parameters a and b are constant coefficients and suggested to be in the range of 1.57 to 2.23 and 3.04 to 15.2, respectively.

Shenton (1985) elaborated a track degradation model based on ballast settlement (Equation 2). In this research, it was noted that ballast deterioration is a factor affecting rail track degradation.

$$S = K_s \frac{A_e}{20}\left((0.69 + 0.028L)N^{0.2} + 2.7 \times 10^{-6} N\right) \qquad (2.2)$$

Where, A_e denotes passing tonnage in Million Gross Tonne (MGT), N denotes the total number of passing axles, K_s is a factor corresponding to the type and size of the sleeper, ballast type and the condition of the subgrade, and L denotes the lift given by tamping machines. It must be mentioned that this research suffers from some drawbacks. For example, a reliable model for measuring the K_s parameter is not defined.

Sato (1995) evaluated track deterioration due to ballast settlement under repeated loading passage. It was noted that for the establishment of a track deterioration model, historical information on the track from its early ages is necessary. For this study, a railway line in Japan was observed. The researcher provided the following equation to estimate the settlement of tamped tracks under frequent loading by train passage:

$$y = \gamma(1 - e^{-\alpha x} + \beta x) \qquad (2.3)$$

Where, y denotes the ballast settlement (mm), x denotes the travel frequency or passed tonnage (million tonnes per year), α, β and γ are coefficients. The first term of this equation is associated with immediate rapid settlement and expresses the process of consolidation/compaction of the gaps between ballast materials. This process is short and should be finished quickly. The second term refers to the linear settlement of rail ballast, which is related to the activity of ballast underneath sleepers.

2.2.2. Statistical Models

A statistical model is a type of mathematical model which can deal with a large amount of data. To establish a statistical model, sufficient historical data are required. Statistical models can be employed to cope with a large number of descriptive factors that can affect rail track degradation. Statistical models can be classified into three main groups: deterministic models, stochastic models and probabilistic models. Each model has sub-categories which are discussed in this section (Soleimanmeigouni, Ahmadi & Kumar 2018).

2.2.2.1. Deterministic Models

A deterministic model is a type of statistical model in which randomness is not involved in the construction of the future condition of the system. Therefore, the model always generates the same output for a given starting condition or state (Baldi et al. 2016). There are different approaches that can be categorised under the deterministic model category. Regression models, classification models and clustering models are among the most applied deterministic models that have been used in rail track degradation studies. In this section different successful examples of deterministic models are investigated and summarised.

Regression models can be categorised into four different types: linear regression, exponential regression, multi-stage and multivariate regression. In this section, the relevant literature on these models is discussed.

Linear regression is one of the simplest statistical techniques for estimating the relationship between a dependent variable and one or a number of independent variables. The best-fitting line represents the relationship between dependent and independent variables. Montgomery et al. (2012) and Westgeest et al. (2012) conducted a linear regression model to predict the effective contributors to track deterioration progress and the volume of maintenance required over a long period of time. In this research, the Key Performance Indicator (KPI) of the track quality was defined as the dependent variable. The KPI is calculated based on the combination of track geometry parameters with their justified coefficients. Different parameters, including the type of tamping, passing tonnage, sleeper type and closeness to switches, were considered as independent variables. According to the analysis, segments with switches have degradation rates faster than others and segments containing concrete sleepers degrade more slowly than segments with hardwood sleepers.

Exponential regression is a type of non-linear regression estimation which can produce the best fit for a set of data. Sadeghi and Askarinejad (2010) developed an exponential regression model to determine rail track degradation rates. In their study, changes in the Track Structure Index (TSI) and the Track Geometry Index (TGI) were considered as dependent parameters. The TSI is mostly based on the condition of rails, sleepers and ballast, while the TGI is mostly based on the condition of twist, alignment, gauge and cross-levels. They used passing tonnage in MGT, time period, initial TSI or initial TGI and the average running speed as the independent variables in their degradation model. Based on analyses of the test zone, two equations for predicting future TGI and TSI were developed. A comparison of the results demonstrated that the sensitivity of TGI to the independent parameters is larger than that of TSI.

Multi-stage regression is a type of linear regression model which has the capability to cope with different stages of degradation prediction. Ahac and Lakušić (2017) developed a tram maintenance-planning framework based on regression models. The models used data from the Zagreb tram network. The gauge deviation value was considered as the dependent variable and the passing traffic, tram speed and the total number of exploitation days were considered as the independent variables. According to the results of the gauge deviation model stiffer rail fastening system and higher track curvature can increase the rate of tram track degradation.

Multivariate statistics is a sub-division of statistical analysis that can analyse more than one dependent variable. Jovanovic et al. (2011) developed a multivariate statistical model for the prediction of railway track geometry deterioration. In this research, a track section in Turkey was observed. Sleeper type, speed, curvature, rail length and the history of maintenance activities (e.g. tamping, rail welding and sleeper renewal) were considered as independent variables, and track geometry parameters were considered as dependent variables. Based on the results of this study, it was determined that rail length had an effect on the deterioration rate with a negative sign. In addition, when the maintenance activities decreased, the renewal activities increased.

A classification model like a supervised learning model is suitable for predicting and describing datasets with nominal or binary categories. In rail track degradation prediction modelling, different types of classification models such as decision trees and logistic regression have been used. A decision tree is a form of supervised classification learning employed to solve binary classification problems in the data mining field. This model provides flexibility in handling a wide variety of input information (numeric, nominal and textual). The outcomes of decision trees can be summarized in a number of logical if-then conditions (Zhu & Taher 2015). Alemazkoor et al. (2015) proposed a decision tree model to predict track geometry degradation. Independent variables were passing tonnage, track length and time gap (elapsed days between two consecutive defect records), and dependent variables were track geometry parameters, including longitudinal level and alignment. According to the results, the prediction accuracy of the decision tree for the output variables was satisfactory.

Nunez et al. (2014) conducted a research to improve the maintenance decision-making process, based on big data from Dutch Railways. In this study, monitoring data for establishing a decision tree model were collected from the Axle Box Acceleration (ABA) measurements. The inspection data were entered in the decision tree algorithm and squats were classified into different categories, including non-problematic, high-density sections, potential severe defects and severe defects. Based on the results of the case study, the application of the model found 100% of larger squats and 85% of small squats. The proposed model can also provide a significant reduction in railway maintenance costs.

Logistic regression is a type of regression model that, due to its capability in classifying data, is considered as a classifier model in rail track degradation prediction modelling (Montgomery, Peck, & Vining 2012; Jr. Hosmer, Lemeshow & Sturdivant 2013; Fagerland & Hosmer 2017).

Hajibabai et al. (2012) used a logistic regression model to forecast the probability of high-impact wheel train stops. In this research, data collected from Wheel

Profile Detectors (WPDs) and Wheel Impact Load Detectors (WILD) were examined by comparing historical measurement records regarding failed and non-failed wheels on the same truck. WILD data (such as vertical average weight, vertical peak force and lateral average force) and WPD data (such as vertical flange, rim thickness and wheel angle) were used in the development of regression models. Based on the results of this study, the regression model developed for the WPD model is not as effective and accurate as of the WILD model in terms of failure prediction.

Andrade and Teixeira and (2014) examined unplanned maintenance needs regarding rail track geometry degradation using a logistic regression model. This study aimed to find the probability of unplanned (corrective) maintenance which must be applied to a given rail track segment. For this purpose, data related to the location of bridges, switches and stations on a Portuguese railway line were gathered. The probability calculation can result in one of two possible values: 1 demonstrates that at least one failure was detected (i.e. unplanned maintenance is required), and 0 demonstrates that no failure was detected. One of the main findings of this study is that the Standard Deviations (SDs) of alignment and longitudinal level are useful and reliable predictors both for planned and unplanned maintenance.

Cluster analysis is a segmentation method applied to identify homogenous objects. The main task of these type of models is to group a set of data into different clusters such that objects in the same cluster are more similar to each other than those assigned to other clusters (Sarstedt & Mooi 2019).

Nicodeme et al. (2017) developed a clustering analysis to monitor the rail surface. This research aimed to propose a new non-destructive method of rail inspection. For this experiment, a multispectral camera mounted on an autonomous train has been used. In this research, unsupervised hierarchical clustering was applied on multispectral images for segmentation and supervised learning applied for recognition. Based on this experimental research, once pollutants or damages are recognised, proper actions will be taken according to the field measurement database. The research proposed that for improving the clustering analysis, using a higher resolution camera can be helpful. In addition, more complex clustering algorithm to process a wide range of parameters and a huge amount of data is recommended.

Jovanović et al. (2015) used a clustering model for their study of rail track degradation in relation to the development of Railway Infrastructure Maintenance Management Systems (RMMSs). The proposed RMMS included different tasks, such as railway condition monitoring, degradation prediction modelling and the development of maintenance and planning strategies. In condition monitoring analysis, a clustering analysis was used to separate rail defects (internal and surface) and classifies defects based on their exceedance of predefined thresholds. This study concluded that to improve the performance of condition monitoring of a railway system, various parameters must be entered into the system regularly, such as the rail type, ballast type, sub-grade condition, fastening condition, sleeper condition and rail corrugation.

2.2.2.2. Probabilistic Models

Probabilistic modelling is a branch of statistic modelling which aims to predict the condition of a system in future. In rail track degradation modelling, probabilistic models employ a distribution pattern to represent the probability of a system's component rate of failure or disruption in a time interval (Gorjian et al. 2010; Jeong et al. 2019). There are different approaches of the probabilistic model which have been applied in rail track degradation prediction modelling such as continuous probability distribution models, Markovian models and Bayesian models. In this section successful application of probabilistic models in rail track degradation prediction have been investigated and summarised.

Various distributions can be used in the modelling of rail track degradation, such as normal distribution, Weibull distribution, Gamma distribution, Gaussian distribution and Dugum distribution.

Caetano and Teixeira (2015) developed an optimisation model for scheduling railway track maintenance and renewal operations. In this research, degradation of different rail track infrastructures such as ballast, sleeper and rail have been studied in order to minimise the track life-cycle cost. In this research, for predicting rail fatigue defects, Weibull law as a probability distribution has been used. The Weibull distribution parameters were forecasted by using the outcomes from the previous studies. For the case study, historical data related to a rail line in Portugal were used. Based on the results of this study, optimal maintenance can be achieved by selecting appropriate time intervals for renewal interventions.

Sadeghi (2010) developed a normal distribution of track geometry data in order to introduce new track geometry indices. Based on the field investigation, the dataset of this study was developed. Four track geometry parameters, gauge, twist, longitudinal level and alignment, were included in this study. By assigning the justified coefficient to each geometry parameter and each combination of them, new indices were defined. The proposed indices were calculated separately according to different track classes. This study concluded that these indexes can be applied to the evaluation of track geometry conditions and maintenance activities.

Audley and Andrews (2013) studied the effects of maintenance on railway track geometry deterioration. In this research, data from the UK Network Rail were examined, and the SD of the longitudinal level was chosen as the dependent variable. The line speed and maintenance history were both independent variables in this research. This paper describes the application of two-parameter Weibull distribution to analyse the distribution of track geometry degradation over the time following maintenance. The results of this study proved the theory that tamping can damage the ballast and causes track geometry to deteriorate more rapidly.

In a Markov model, the fundamental assumption is that the probabilities of transferring from a state to any other state rely only on the current state, and not on the procedure by which the current state is reached. This characteristic is called the memory-less property of the Markov model. More details can be found in (Zhang, Kim & Tee 2017).

Yousefikia et al. (2014) developed a Markov model for the prediction of track deterioration in Melbourne trams in an attempt to determine optimal maintenance planning. In their studies, the effect of horizontal tight curves on rail wear was considered. They defined three states of track condition according to tram safety operations. The first degradation state represents minor degraded failure. When the second degradation state is identified, immediate maintenance is required. Finally, operational restrictions are required when the final state is identified. By applying a Markov model, transitions from different states were established. This study emphasised that by applying preventive maintenance, the transition to the last state or operational restrictions can be avoided.

Bai et al. (2015) developed a Markov chain model to predict rail track irregularities based on Chinese railway maintenance management data. In this research Track Quality Index which represents the quality and condition of rail track sections has been categorised into four states. A Markov stochastic process has been used to create a transition matrix of degradation. The transition matrix represents a degradation from a specific state to another state because of irregularities in rail track sections. For verification of the proposed model, data measured by a track geometry car has been used. The results from the assessment approved the model performance but it has been noted that for improving the accuracy of the model, a large amount of historical data are required.

Sharma et al. (2018) developed a Markovian prediction model based on Track Quality Index (TQI) in order to predict the occurrence of geometry defects in heavy rails tracks. Different factors have been included in the above studies to predict the deviation of rail track geometry parameters and failure rate. Daily traffic (in MGT), the total number of exploitation days and previous geometry and structural deviation measurements are among the important variables.

The integration of prior information and data is handled by Bayesian rules which provide a probabilistic mechanism for learning from data (Yousefikia et al. 2014).

Andrade and Teixeira (2012) developed a Bayesian model line to assess rail track geometry degradation. In this research, the SD of the longitudinal level was considered as the dependent parameter and primary SD longitudinal level measured after tamping operations or renewal, the rate of deterioration and the accumulated tonnage since tamping operations or renewal (in MGTs) were the dependent variables. The case study was sourced from a Portugal rail network. Log-normal prior distribution was applied and Markov Chain Monte Carlo (MCMC) simulation was performed to obtain the fitting parameters of the distribution. This study concluded that mid-term and long-term maintenance and renewal plans must be undertaken after 2-3 years.

Andrade and Teixeira (2015) developed hierarchical Bayesian models for the study of rail track geometry degradation. This study aimed to monitor the evolution of the SD of alignment and the SD of longitudinal level as two important quality indicators associated with railway track geometry. A section of the Lisbon–Oporto line was considered as a case study in this research, and posterior distribution was used to represent the data. An MCMC model was used to solve the Bayesian model. According to the results of the case study, the accuracy of the prediction of the SD of alignment is limited compared with that of the SD of longitudinal level parameters.

Jamshidi et al. (2016) developed a probabilistic defect-based risk assessment model for rail failures in railway infrastructure based on a Bayesian model. This research addressed the deterioration of rails because of squat (a type of defect related to rail surface) growth. An exponential correlation between the visual length of a squat and squat crack depth was investigated in an attempt to analyse the severity categories. In this research, a non-linear regression model was developed and posterior distribution was used. An MCMC model was applied to find the fitting parameters. The failure risk factor discussed in this research can be used to represent the health status of rails and maintenance planning.

2.2.2.3. Stochastic Models

A stochastic model is a statistical model which contains one or more random variables in rail track degradation. Uncertainty, an inherent characteristic of infrastructure deterioration, is captured in this type of model (Baldi et al. 2016). Different approaches can be categorised under the stochastic models such as stochastic probability distribution, time series, Petri Nets (PNs) and survival models. In this section different examples of these approaches are investigated and summarised.

Mercier et al. (2012) developed a bivariate Gamma process model for track geometry intervention schedule. The main purpose of the research was to discuss intervention scheduling and its impacts on rail track performance. In this research, as the SD of longitudinal level and alignment were considered as degradation indicators, the application of a bivariate model was inevitable. As a case study, the Paris-Lyon high-speed line was selected. According to the results of this study, the maintenance scheduling derived from the two proposed degradation indicators was much more reliable than those based on a single indicator.

Vale and Lurdes (2013) proposed a probabilistic model for the prediction of rail track geometry degradation on a Portuguese railway line. In this study, the SD of the longitudinal level was considered as the main dependent variable. The SD of the longitudinal level with respect to three-speed ranges (based on the maximum allowed train speed) and the rail position (right and left) were involved in this study. For the probabilistic analysis, the Dagum distribution was used. The application of the Dagum distribution to geometrical rail track degradation was considered as the major contribution of this study.

Zhu et al. (2013) employed a Gaussian random process model for the estimation of changes in track irregularities. The track irregularity data used in this paper were collected from a railway line in China, although only alignment and longitudinal level were used. As the data provided in this study were of a Gaussian nature, Power Spectrum Density (PSD) and level crossing estimation were applied. Based on the results of this research, the applications of these models on track irregularities is useful to enhance the evaluation of railway track conditions.

A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally-spaced points in time. Therefore, it is a sequence of discrete-time data.

Quiroga and Schineder (2010) developed an auto-regressive model to forecast railway track geometry deterioration. The Auto-Regressive Moving Average (ARMA) model is another type of time series problem-solving model. The SD of the longitudinal level was considered as an effective parameter to indicate rail track geometry degradation. Previous longitudinal level values, the total length of section and length of tamped tracks were used to predict the indicator parameter. They applied this model to a section of a French high-speed railway line. The study concluded that the proposed model can be used by tamping scheduling optimization systems.

Jia et al. (2012) used a Kalman filtering model for solving the track irregularity time series. In this research, a Kalman filter was used to forecast the values of cross-level parameters. Kalman filtering can be employed to predict the current state when the estimated state from the last time and the current state are known, regardless of consideration of estimates or historically informative observations. In this research, passing tonnage, track geometry data and train speed were used as indicators for the prediction of track state. This study concluded that the proposed model was successful in the prediction of future cross-level parameters.

Salvador et al. (2016) applied time series models for railway track monitoring purposes. In this research, time-frequency characterisation associated with railway tracks were analysed to fulfil the demand for high quality and cost-effective maintenance actions. For the case study, a rail line of the Metropolitan Rail Network of Valencia (Spain) has been used. In this study, axle box acceleration, train speed and track geometry parameters have been applied as the main parameter to detect rail track irregularities. The results of this study implied that wide ranges of vibration modes measured by axle box accelerometers along with the train speed can be used to monitor the condition of the rail track infrastructure.

PNs are useful graphical-mathematical models consisting of transitions, places (states) and arcs. Petri nets provide a graphical notation for performing stepwise processes that contain choice, condition, iteration and concurrent execution. Petri nets models have the ability to model the combination of degradation and maintenance (Rana, Verma & Srividya 2016). Prescott and Andrews (2013) proposed a PN model for track ballast maintenance inspection. As the PN consisted of different places and transitions, the proposed model covered different states of maintenance, including good conditions, normal maintenance and traffic closures. In this research, geometry parameters, including longitudinal level, alignment, gauge and twist, were used to determine the places of the PN model. This study emphasised that to establish a PN model, maintenance history is an important factor to prioritise future maintenance activities. In this regard, track sections with severe levels of ballast deterioration have higher priority.

Andrews et al. (2014) developed a PN model to predict railway track deterioration. The proposed model aimed to incorporate track deterioration and all activities related to repair and renewal. The SD of the longitudinal level was employed to indicate the track geometry quality. Rail type, sleeper type and line traffic were included in the development of this model. A two-parameter Weibull distribution was used to demonstrate the distribution of times to track deterioration. The data employed in this research were sourced from a section of a UK railway. A series of recommendations were made. For instance, the average routine repair time of 50 days can often keep track in excellent and reliable condition.

Survival model is a branch of stochastic models, analyse the expected duration of time before the occurrence of one or more events such as failure or severe restriction in mechanical systems. Survival analysis tries to find what proportion of a population will survive after a certain time or at what rate they will fail or cease operation. He et al. (2013) developed a survival model to assess risk and an optimization model for repair decisions, in an attempt to reduce the probability of train derailment. In railway maintenance, if no derailment occurs between two scheduled inspections on a track segment, the track is considered a survived track; otherwise, the track segment is considered a failed track. The dataset used in this research covered 3-year defect data and derailment data from a railway line in the United States. The predictor variables used in this research were monthly traffic in MGTs, the number of geometry defects and 90 percentile amplitude of geometry defects. The Cox-Snell residual was used to evaluate the proposed model. According to the results of the model, all track geometry parameters and rail wear had positive impacts on derailment risk.

Moridpour and Hesami (2015) conducted a research to estimate the degradation and performance of Melbourne tram tracks. The main objective of this research was to investigate maintenance needs associated with tram track geometry defects based on survival models. The dataset of this study was collected from the Melbourne tram network. The probability of reaching the maintenance limit for different fault classifications (speed restriction, maintenance intervention and traffic restriction) and track categories was estimated. According to the results of the model, curved type and H-crossing tracks have the highest likelihood of failure compared to other types of track. In other words, as the survival probability for curve tracks is lower at all times, these tracks are more prone to reaching the maintenance limit.

2.2.3. Artificial Intelligence (AI) Models

In recent years, AI-based models have become popular, as they overcome the deficiencies of current mechanistic models in the prediction of rail track degradation. AI models involve activities and developments relating to human-like intelligence reproduced by computer applications. For this purpose, they exploit computer techniques or reasoning algorithms that attempt to automate intelligent functions (Jovanović, Guler & Čoko 2015). AI models can be categorised into main sub-categories, including Artificial Neural Networks (ANNs), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Decision Support Systems (DSSs) and machine learning model. In this section, the application of these models in rail track degradation is discussed.

2.2.3.1 ANNs

ANNs consist of a number of neurons and these neurons make communication with each other through weighted connections. Neurons in the neural network are connected together in a multi-layer structure. The output of each neuron is transferred to the next neuron through a connection and is the input for that neuron (Yadav, Yadav & Kumar 2015; Guler 2013). Different examples of the implementation of this type of model are presented in this section.

Sadeghi and Askarinejad (2012) proposed an ANN model for the evaluation of railway track quality. This study investigated the possibility of relationships between track geometry defects and track structural problems. A multilayer feed-forward network was used as the architecture of the network. The network input was the SDs of track geometry data (gauge, longitudinal level, alignment and twist) and the output presented the predicted defect density of track structural components (sleeper, rail, ballast and fastening). Defect density was calculated by dividing the number of damaged units (a unit is a distance between two successive sleepers) by the total number of units included in a track segment. Based on the results of this study, the proposed ANN model had better accuracy for low and medium quality track conditions than new or high-quality tracks.

Moridpour et al. (2017) elaborated an ANN model for predicting the tram track degradation using track maintenance data and addressing the curved sections only. In this research, the Melbourne tram network was used as a case study. A multilayer feed-forward ANN model with three layers was applied in the research to predict the dependent variable. Different variables such as rail type, rail profile, passing tonnage in MGT and the instalment year have been included to predict the deviation of track gauge parameter. Based on the results of this study, the type of tracks and last gauge measurement, have a significant impact on the track geometry deviation. The developed model presented a reasonably good prediction accuracy.

2.2.3.2. ANFIS

The combination of ANN and a Fuzzy Inference Engine (FIS) is called an ANFIS model. Since it integrates the principles of both fuzzy logic and neural networks, it has the ability to obtain the benefits of both in a single framework (Zimmermann 2010). In this section, different examples of this model are presented.

Dell'Orco et al. (2008) developed an ANFIS model for optimising rail track maintenance and planning issues. The proposed model included five input parameters and one output parameter. The inputs were the SDs of geometry parameters (including alignment, longitudinal level and cross-level), the number of the days elapsed from the latest tamping and the number of previous tamping works. The output of the system was the number of days passed from the last tamping to the following one. This model was applied to a specific line of Italian Railways. For the purpose of model validation, the root mean square of the differences between the output of training data at each epoch and the output of ANFIS was calculated. It was found that the model could provide correct dates prior to or equal to the maintenance threshold.

Shafahi et al. (2008) proposed an ANFIS model to predict rail track degradation on the Iranian Railway network. The data bank for this study consisted of different types of parameters, such as annual traffic, construction date and the number of passing axles. The CTR index was examined as the main parameter for the track prediction. CTR is applied to evaluate railway tracks in terms of quality and geometric condition. The CTR can vary from 0 to 100 where 100 represents the best possible track condition. In accordance with traffic condition and geographical location parameters, six track classes were organised, and backward propagation was used for training the ANFIS algorithm. The results showed that most of the estimations were similar to real outcomes.

Karimpour et al. (2018) elaborated an ANFIS model to predict rail track degradation based on the gauge parameter. In this study, the dataset of the Melbourne tram network has been used. This study suggested that an accurate model is able to play a significant role in predicting the long-term performance of rail tracks. Gauge deviation parameters associated with the previous year and two years ago were among the main parameters in the model development. The results show that the model can predict the gauge deviation for the coming year with acceptable accuracy.

2.2.3.3. DSS

A Decision Support System (DSS) is a computer application to support experts in decision-making processes by using decision rules and analytical techniques. This type of system is developed to help decision-makers to solve both unstructured and semi-structured problems (Rashidi, Samali & Sharafi 2016). Some examples of implementations of DSSs in rail track degradation are provided in this section.

Guler (2013) developed a DSS application to perform railway track maintenance and renewal management activities. The dataset of this study was sourced from the Turkish State Railway. Different parameters were covered by this system, including the type of ballast, tamping history, gauge value, number of trains, age of rails, speeds and cost analysis. Various maintenance and renewal operation were addressed in this study, such as ballast renewal, rail renewal and rail lubrication. Operation actions introduced in this system were classified into four categories: do nothing, regular maintenance and renewal actions, corrective maintenance and finally traffic prohibition. The results of the case study showed that the system decided in a reliable manner and system performance could be enhanced by adding new rules and more calibration limit values.

Morant et al. (2016) developed a model-oriented decision support system for the maintenance of railway signalling systems. The reliability of signalling systems directly affects the availability of railway networks. This system included various corrective maintenance parameters. The proposed DSS was based on failure analysis of signalling systems and corrective maintenance interventions carried out in the past. The proposed system can create preventive maintenance policies and strategies depending on different signalling failures. As a case study, a rail line with few changes in many years has been selected. Historical maintenance data related to the line have been analysed and entered into the system. Based on the results of the system assessment, the implementation of the proposed system was successful. As the proposed model was highly dependent on empirical data, the authors recommended a large amount of information for reducing the system's limitations.

2.2.3.4. Machine Learning

Machine learning is a branch of AI which can be applied to both historical and real-time data to forecast the future conditions of a system. Machine learning models are employed in tasks where due to uncertainty, the design and development of explicit algorithms are impractical. Support Vector Machine (SVM) and Random Forest (RF) is a branch of machine learning which has the ability to establish prediction models based on both categorical and numerical output variables (Michalski, Carbonell & Mitchell 2013). In this section, different examples of the implementation of SVM and RF models in railway condition monitoring are described.

Asada et al. (2013) developed an algorithm for railway condition monitoring and fault detection based on an SVM model and designed a fault detection and diagnosis system for point machines. A point machine is an actuator which drives the switch blade from one position to the opposite position in order to offer different routes to trains. In this research, two faults were investigated: under-driving and over-driving of the drive rod. Data from a Japanese AC point machine were collected. The results of system validation demonstrated that the proposed model had accurate performance in fault detection by applying it to the electrically active power data of the point machine. Moreover, it could be applied to other similar infrastructure, such as level crossings and train doors.

Li et al. (2014) reported machine-learning models to forecast impending defects and alarms of critical components of rail cars. In this study, learned rules were developed based on historical data to predict which rail cars were likely to have problems and to predict intensive existing alarms prior to a real alarm event in an attempt to decrease instant train stops. The model development included five steps: feature extraction, dimension reduction, model training, prediction and confidence estimation and rule simplification. For the evaluation, the results of the proposed SVM model and a decision tree were compared against the same data. Based on the results, the customized SVM model showed better performance than the decision tree for alarm prediction.

Lasisi and Attoh-Okin (2018) developed a machine learning approach based on the SVM model to predict deviation from the pre-defined TQI threshold. The proposed model was intended to apply in maintenance planning and scheduling systems. In this study, track geometry parameters along with other important rail parameters such as velocity, surface, rail profile and traffic volume were considered. Dataset was gathered from a US Class I railroad. In this study, track sections were divided into specific lengths for calculating TQI values. The performance of the model was assessed by applying the True Positive Rate (TPR) and False Positive Rate (FPR). The validation showed that the proposed SVM models were able to classify successfully track geometry defects based on the TQI.

Falamarzi et al. (2018b) developed a Random Forests (RF) model to predict the future deterioration index. In this study, the Melbourne tram network has been used as the case study and gauge deviation parameter is selected as the main parameter to develop the index. Based on the results of this research, the adjusted R^2 value of the proposed model is considerably high and the prediction error is negligible, which demonstrates that the model has the reasonable performance in predicting the deterioration index.

2.2.4. Summary of Degradation Models

In this section, various examples of mechanistic railway degradation models, statistical models and AI models were reviewed. The main disadvantage of conventional mechanistic models is their inability to deal with the inherent uncertainty of track degradation behaviour. Inability to deal with inhomogeneous track sections and also consider different variables involved in track degradation are the main limitations of mechanistic models. However, this limitation can be resolved by applying empirical mechanistic models. Furthermore, different type of statistical models was examined. Ability to work with a large amount of data and deal with the inherent uncertainty of degradation parameters are among their strengths. Difficulties to determine the probability distribution of probabilistic models and the potential to ignore the degradation factors in deterministic models are among the major limitations of statistical models.

Lastly, AI models have been discussed. Calibrating degradation models by optimising effective parameters and emulating the decision-making process of human experts are among the advantages of AI models. The lack of transparency on how decision and outcomes are reached is the main disadvantages of AI models. Based on the research outlined above, various parameters are involved in rail degradation models. In the mechanistic model ballast settlement is a significant factor in mechanistic degradation models. In statistical and AI models, tamping, the interval of maintenance activities, deviation of track geometry parameters measured in previous years, SD of track geometry parameters, passing traffic, train speed, traffic type (type of trains) and traffic density are among the main degradation parameters. The main degradation parameters in rail track analysis are listed in Table 2.1.

Table 2.1: List of degradation models and applied variables

Models	**Independent variable**	**Dependent variables**
• Mechanistic models • Statistical models ✓ Deterministic ✓ Probabilistic ✓ Stochastic • AI models ✓ ANN ✓ ANFIS ✓ DSS ✓ Machine learning	• Maintenance history • Days from last maintenance • Days from the last tamping • Type of tamping • Time • Exploitation days • Rail age • Days between two consecutive defects • Ballast /sub-ballast conditions • Ballast settlement/pressure • Track geo. parameters • Initial TGI • Track stru. parameters • Initial TSI • Sleeper type • Fastening system • Closeness to bridges/switches • Expansion joint condition • Rail condition/length • Wear • Age • Internal defects • Corrugation • Wheel condition • Wheel defects • Angle of attack • Traffic condition • Train (maximum) speed • Posted speed • Monthly passing tonnage • Accumulated tonnage • Repeated loading number • Environmental condition • Temperature • Soil type • Geographical condition • Mountain/hilly • Curve radius • Acceleration data	• Ballast settlement rate • TGI index • TSI index • Rate of wear • Squat rate • Failed wheels • Ballast degradation • Maintenance limit • Derailment risk • Tamping remains days • Maintenance activities • Alarm prediction • Vertical acceleration • Point machine failure • Future track geometry parameters • Gauge • Twist • Longitudinal level • Alignment • Cross-level • Future track structural parameters • Condition of sleepers • Condition of the fastening system • Condition of expansion joints

2.3. Track Degradation Indices

Various studies have been conducted in the field of railway track degradation modelling. However, few studies have attempted to develop and design track degradation indices. Track degradation indices as track quality representatives are important as they are mostly created based on data collected over several years. In this section, different track degradation indices based on track geometry parameters used in different researches are discussed. Track geometry parameters can be categorised into the gauge, twist, alignment, profile and cross-level (Figure 2.2).

Track gauge is the right-angle distance between two rails at a given location, below the top surface of the railhead. A gauge defect is a deviation from the prescribed value. Cross-level is the difference between the top surfaces of two rails at a specific location Track twist is the numerical difference between two cross-levels measured at a predefined distance of a rail track apart. Profile is the change in elevation in a specific chord length. Alignment is the difference between the actual horizontal alignment and the designated alignment (He et al. 2015).

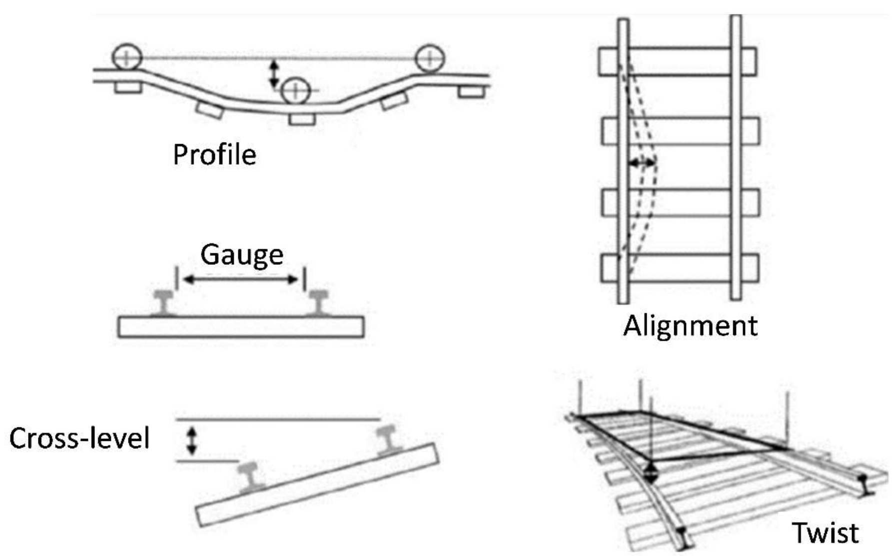

Figure 2.2: Track geometry parameters (Giacomo et al. 2018)

2.3.1. Existing Degradation Indices

Different approaches and mathematical equations have been applied to formulate track degradation indices. For instance, a synthetic track quality coefficient is used in Poland to evaluate the track geometry condition based on the Standard Deviations (SDs) of different track geometry parameters. The proposed index is expressed by (Madejski & Grabczyk 2002; Chudzikiewicz et al. 2017):

$$J = \frac{S_z + S_y + S_w + 0.5S_e}{3.5} \tag{2.4}$$

where J is the proposed track degradation index, S_z represents the SD of profile, S_y denotes the SD of alignment, S_w is the SD of track twist and S_e represents the SD of track gauge. The chord length (representing the length of measurement for collecting track geometry parameters) of 10 metres is used in this study. The permissible values for the Polish index based on train speed are presented in Table 2.2.

Table 2.2: Allowable values of *J* index for different speeds.

Speed (km/h)	**30**	**40**	**90**	**120**	**160**	**200**
J index	12.0	11.0	6.2	4.0	2.0	1.4

In the USA, a track roughness index was established by Amtrak to represent the condition of rail tracks. This index can be calculated by the average of squared differential geometry deviations over a chord length of 20 metres as follows (Westhuizen & Grabe 2013; Liu et al. 2015):

$$r^2 = \frac{1}{n}\sum_{i=1}^{n-1}(G_{dev_{t+1}} - G_{dev_t})^2 \tag{2.5}$$

Where r^2 represents track roughness value, n is the number of measurements and $G_{dev_{t+1}}$ and G_{dev_t} represent the amount of gauge deviation for two consecutive years. The proposed index can be utilised for gauge, cross-level, alignment and profile. Table 2.3 demonstrates the condition of tracks associated with the roughness index.

Table 2.3: Condition of track in accordance with the roughness index.

r^2 **value**	$r^2 < 1$	$1 < r^2 < 4$	$4 < r^2 < 16$	$r^2 < 16$
Track condition	Very good	Good	Average	Poor

The Canadian National Railway Company (CN) has introduced a Track Quality Index (TQI) based on the squared SDs of track geometry parameters as follows (Setiawan & Rosyidi 2016; Roghani 2017):

$$TQI_i = 1000 - C \times \sigma_i^2 \tag{2.6}$$

Where TQI represents the proposed track quality index, C is constant and is determined as 700 for main track lines and σ_i is the SD of track geometry parameters (gauge, cross-level, alignment and profile). The above equation can be used for assessing individual track geometry parameters. In order to calculate overall TQI, the average of indices for track geometry parameters has been proposed by CN. Accordingly, a smaller TQI indicates that the track segments are at a greater risk of failure.

In China, the Chinese national railroads use the sum of SDs of total track geometry parameters for calculating the overall TQI as follows (Xu et al. 2011; Li et al. 2016):

$$TQI = \sum_{i=1}^{n} \sigma_i \tag{2.7}$$

Where TQI represents the track quality index and σ_i is the SD of track geometry parameters. In this research, two different lengths for assessing overall track quality are proposed. For high-speed railroads, a track length of 500 metres is applied and for conventional railroads, a track length of 200 meters is applied. Larger TQI values imply a potential reduction in overall track quality.

Railway track geometry parameters can be assessed according to European Standard EN 13848. This standard can be applied to profile, alignment and gauge parameters. To calculate this index, the SDs of the alignment and profile are used. For the gauge parameter, the difference between the mean value of the gauge and the pre-defined gauge value is applied. In Table 2.4, recommendations of EN 13848 for profiles based on Track Quality Classes (TQCs) and two specific speed categories are shown. The TQC "A" represents the safest condition of a rail track and the TQC "E" represents the critical condition of a rail track. In Table 2.5, the thresholds for assessment of gauge parameter based on the different values and two specific speed categories over a length of a 100 m track segment are tabulated (Berawi et al. 2010).

Table 2.4: Recommendations of EN 13848 for assessment of profile.

Speed (km/h)	**Threshold of SD (mm)**				
	TQC				
	A	**B**	**C**	**D**	**E**
80<V≤120	<0.75	1.10	1.80	2.50	>2.50
120<V≤160	<0.65	0.85	1.40	1.85	>1.85

Table 2.5: Recommendations of EN 13848 for assessment of gauge.

Speed (km/h)	**Difference between the mean value of gauge and pre-defined gauge value**		
	Alert Limit (SL)	**Intervention Limit (IL)**	**Safety Limit (AL)**
80<V≤120	-7≤ difference ≤27	-6≤ difference≤25	-5≤ difference ≤22
120<V≤160	-5≤ difference ≤20	-4≤ difference ≤18	-3≤ difference ≤16

In India, the national railway uses an exponential equation to calculate the TQI. The index is defined for individual track geometry parameters (e.g. gauge and profile) as follows (Mundrey 2009; Berawi 2013):

$$GI = 100 \times e^{-(SD_m - SD_N)/(SD_U - SD_N)} \tag{2.8}$$

where *GI* is the individual track geometry index, SD_m represents the SD of the measured parameter, SD_U denotes the SD value of a track with immediate maintenance need, and SD_N indicates the SD of a newly-laid track. The values of SD_U and SD_N are constant and can be found in Table 2.6.

Table 2.6: The values of SD_N and SD_U based on the track chord length and speed.

Parameters	**Segment length (m)**	**SD_N (mm)**	**SD_U**	
			Speed > 105 km/h	**Speed < 105 km/h**
Alignment	7.2	1.5	3	3
Gauge	-	1	3.6	3.6
Twist	3.6	1.75	3.8	3.8
Unevenness	9.6	2.5	6.2	7.2

The overall Track Geometry Index (TGI) can be obtained from:

$$TGI = \frac{2UI + TI + 6AI + GI}{10} \tag{2.9}$$

where TGI represents the overall track geometry index, UI is the index for unevenness, TI indicates the index for twist, AI denotes the index for alignment and GI is the index for gauge, which can be calculated using Equation 2.8. The allowable amounts of *TGI* are shown in Table 2.7.

Table 2.7: Condition of track based on *TGI* values.

***TGI* value**	***TGI*<36**	**36<*TGI*<50**	**50<*TGI*<80**	**80<*TGI***
Track condition	Poor	Average	Good	Excellent

The Swedish national rail network uses its own TQI to assess track geometry condition. This index is based on the SDs of track parameters (Andrade & Teixeira 2012; Odolinski & Smith 2016) and is obtained from:

$$Q = 150 - \frac{100\left[\frac{\sigma_H}{\sigma_{H_{lim}}} + \frac{\sigma_S}{\sigma_{S_{lim}}}\right]}{3} \quad (2.10)$$

where Q represents the index for evaluating track geometry condition, σ_H is the average of SDs of left and right profiles, σ_H denotes the SDs of other track geometry parameters including gauge, cross-level and horizontal deviation, $\sigma_{H_{lim}}$constitutes the allowable limit of σ_H, and $\sigma_{S_{lim}}$ indicates the allowable limit of σ_S based on track type. According to Swedish standards, the maximum and minimum values for the proposed index are 50 and 150, respectively. The allowable values of the index are in the range of 70 to 90. The chord length of 12 m is used to measure the deviations.

In Iran, based on SD values and the mean values of the track geometry parameters, *TGI*s for individual parameters and overall TGI (*OTGI*) have been proposed. A chord length of 19 m is used in the research for collecting the deviation values of parameters. These parameters include track gauge, alignment, profile, and twist. The following equation is used for assessing the alignment, gauge and twist parameters (Sadeghi 2010):

$$\mathrm{AI} = \frac{\left|\bar{x}_{AllignLeft}\right| + 3 \times SD_{AllignLeft} + \left|\bar{x}_{AllignRight}\right| + 3 \times SD_{AllignRight}}{2} \quad (2.11)$$

$$\mathrm{GI}^{+} = \left|\bar{x}_{Gauge} + 3 \times SD_{Gauge}\right| \quad \text{Positive Gauge Index}$$
$$\mathrm{GI}^{-} = \left|\bar{x}_{Gauge} - 3 \times SD_{Gauge}\right| \quad \text{Negative Gauge Index} \quad (2.12)$$

$$\mathrm{TI} = \left|\bar{x}_{Twist}\right| + 3 \times SD_{Twist} \quad (2.13)$$

Where AI represents the value of the index for allignment. GI^{+} represents the value of the index for positive gauge. GI^{-} represent the value of the index for negative gauge. TI represents the value of the index for twist parameter.

$\bar{x}_{AllignLeft}$ and $\bar{x}_{AllignRigh}$ represent the mean value of alignment for left and right rails. $SD_{AllignLeft}$ and $SD_{AllignRigh}$ represent the SD of alignment for left and right rails. $\bar{x}_{Gauge}$ represent the mean value of the gauge. SD_{Gauge} represent the SD of the gauge. $\bar{x}_{Twist}$ represent the mean value of twist. SD_{Twist} represent the SD of twist.

For overall track geometry assessment, the following formula has been proposed:

$$OTGI = \frac{\frac{a}{2} \times GI^{+} + \frac{\acute{a}}{2} \times GI^{-} + b \times AI + c \times PI + d \times TI}{\frac{a + \acute{a}}{2} + b + c + d} \tag{2.14}$$

Where $OTGI$ represents the overall TGI, GI^{+} is the positive gauge index, GI^{-} indicates the negative gauge index, AI, PI, and TI represent respectively the alignment, the profile and the twist indices, a, $\acute{a}$, b, c, and d are constant parameters which vary between 0.08 and 1.00, based on track class and the number of defects in a certain chord length. Table 2.8 shows the allowable values of *OTGI* for different track classes and the maximum of two defects.

Table 2.8: Allowable values of *OTG* based on track class.

Track class	**Allowable values of *OTG***
A	2.19<*OTGI*<4.91
B	3.02<*OTGI*<6.26
C	3.62<*OTGI*<7.23
D	7.06<*OTGI*<8.70

The Austrian Railway has proposed a track defectiveness index based on the ratio between the total length of segments that have exceeded the acceptable limit and the total length of the track. According to this method, a defectiveness index for each track geometry parameter can be calculated using the following formula (Madejski & Grabczyk 2002):

$$w = \frac{\sum L_i}{L} \tag{2.15}$$

Where w is the defectiveness index of geometry parameters, L_i represents the sum of the length of segments that exceed the acceptable range and L denotes the total length of track segments. Larger values of the track defectiveness index represent a reduction in track quality. The overall track geometry defectiveness index (five-parameter index) can be calculated by:

$$w_5 = 1 - (1 - w_e)(1 - w_g)(1 - w_y)(1 - w_z)(1 - w_w) \tag{2.16}$$

Where w_5 is the five-parameter defectiveness index, w_g represents the defectiveness of cross-level, w_e is the defectiveness of gauge and w_y denote the arithmetic average of profile, w_z and w_w respectively denote the arithmetic average of alignment and twist. The Austrian Railway has provided the following table for track condition assessment based on overall defectiveness index:

Table 2.9: Condition of track based on w_5.

Track condition	**w_5 values**
New tracks	w_5<0.1
Good condition	0.1< w_5<0.2
Average condition	0.1< w_5<0.6
Poor condition	w_5>0.6

2.3.2. Summary of Track Degradation Indices

As described in the above section, different formulation has been used to develop track degradation indices. Mean value of track geometry parameter and differential geometry deviation or SD of track geometry parameters are among the most applied factors used to formulate the track degradation indices. The mean value of the geometry deviations is considered as an essential factor in the formulation of the track degradation index, as higher values represent larger deviations from the predefined geometry parameter and eventually more risk of track degradation compared to the ones with the lower mean values. Furthermore, the role of the average differential geometry deviation is important. It is notable that two different track segments can have an accidentally equal mean value of geometry deviation, but the one with a higher value of differential geometry deviation reflects a faster rate of degradation in track geometry compared to the segment with a lower differential geometry deviation. Based on these findings, track degradation indices which neglect and miss one of the above factors cannot deal with the degradation process properly. Also according to the literature review, most studies applied in rail track degradation indices deal with heavy rail while light rail such as tram rail is not well discussed. The other limitations of current track degradation indices are the lack of the validation of their application on different case studies. As a result, there are no studies that have been conducted to compare the performance of the different indices. In the following table, degradation indices, the parameters included in their development and their application have been summarised:

Table 2.10: Summary of the track degradation indices

Track degradation index	**Formulation**	**Parameters**
J Index and Chinese Index	SD of track geometry primates	Gauge, twist, profile, alignment, cross-level
Amtrak	Differential geometry deviation	
Canadian Index	Squared of SD of track geometry primates	
Indian national railway Index	An exponential form of SD of track geometry parameter	
Swedish national rail	SD and allowable SD of track geometry parameter	
Iranian national rail	SD and the mean value of track geometry parameter	
Austrian railway index	Length of defectiveness	
European Standard EN 13848	SD and mean value of track geometry deviation	

3.4. Gap of Knowledge

Investigating the existing literature on rail track degradation prediction models and also the indices for track degradation models reveals that there are some shortcomings. For reaching the aim of the study which is developing a cost-effective approach for predicating tram track degradation, the following gaps have been determined and are tabulated in Table 2.11 as follows:

Table 2.11: list of gaps in knowledge that are founded.

Gaps that need to be covered	**Description**
Studies in the area of tram track degradation	Few studies have been conducted in the field of tram track degradation modelling
Existing tram track degradation models	More elaborate and effective prediction models need to be developed
Track degradation indices for tram track	Current indices mostly focused on heavy rail and development of indices based on tram track systems is neglected
Cost-effective approaches for track degradation prediction	Current approaches for tram degradation require physical track geometry measurements which is costly and time-consuming

2.5. Summary

In this chapter firstly wide ranges of degradation prediction models in rail tracks were investigated and the parameters involved in these models were reviewed. Rail track degradation prediction models are applied by rail maintenance and management systems to predict the future condition of rail tracks. These models are considered as the core of preventive rail track maintenance strategies. Finally, the advantages and limitations of degradation prediction models were examined in the summary. Afterwards, the review of current track degradation indices applied in rail maintenance and prediction models were presented. Track degradation index is a useful measure for infrastructure maintenance management systems as well as prioritising and ranking rail track segments with maintenance needs. Different indices with different formulation and methodologies have been investigated. The allowable limit of the indices was specified and at the end, a summary of the existing indices along with the limitations of current indices was provided. Finally, with regard to the literature investigated in the previous sections, the gaps of knowledge in relation to tram track degradation prediction modelling were provided.

CHAPTER 3

DATASET AND RESEARCH FRAMEWORK

3.1. Introduction

In this chapter, first the case study and the dataset of this research are explained. The Melbourne tram network has been selected as the case study of this research. Dataset of this research which contain different track geometry parameters and track structural parameters are described and the effective geometry parameters in tram track degradation prediction are determined. Afterwards, with regard to the research questions, research objectives, research aim and gap of knowledge presented in the previous chapters, the research framework is introduced and explained in the form of a flowchart. The research framework of the study demonstrates the process and steps that are required to be carried out for achieving the main objectives and finally the aim of the research. At the end, the summary of this chapter is provided.

3.2. Dataset

For the case study, the dataset of the Melbourne tram network, which is the longest tram network in the world, have been used. The first electric tram in Melbourne was built in 1889. This tram system is fed by a pantograph sliding on an overhead line. Melbourne tram system consists of 493 trams, 24 routes, and 1,763 tram stops. Figure 3.1 represents the map of the Melbourne tram network. Melbourne tram network covers 250 kilometres of track and runs 31,500 scheduled tram services every week. According to Yarra Tram which is the main operator and manager of the tram network, the total number of patronage in 2017-2018 was 206.3 million. Melbourne tram patronage over five consecutive years is depicted in Figure 3.2 (PTV 2018).

For collecting track geometry parameters, a non-contact optical laser measurement system which has been mounted on a Track Recording Vehicles (TRV) has been used. The dataset of this study consists of different types of track geometry parameters including gauge, profile, alignment, twist and cross-level. Gauge is the deviation from the pre-defined distance between the inner surfaces of the rails in a rail track. Profile is the vertical deviation in a specific chord length. Alignment is the horizontal deviation from the designated alignment. Cross-level is the deviation between the top of head surfaces of the rails at a specific location. Twist is the difference between two cross-levels measured in a specific chord length.

In addition, train traffic in Million Gross Tonnes (MGT) and rail track structural parameters such as rail profile, track surface, rail support and rail type are included in the dataset. Rail profile is the cross-sectional shape of a tram rail, which is presented by kilogram per metre. The track surface is the material of pavement laid down between the tracks and categorised into the asphalt and concrete surfaces. Rail support is rail ties laid perpendicular to rail and categorised into steel sleepers, timber sleepers and concrete sleepers. Rail type is the shape of tram railhead and categorised Grooved and T-shapes (PTV 2018).

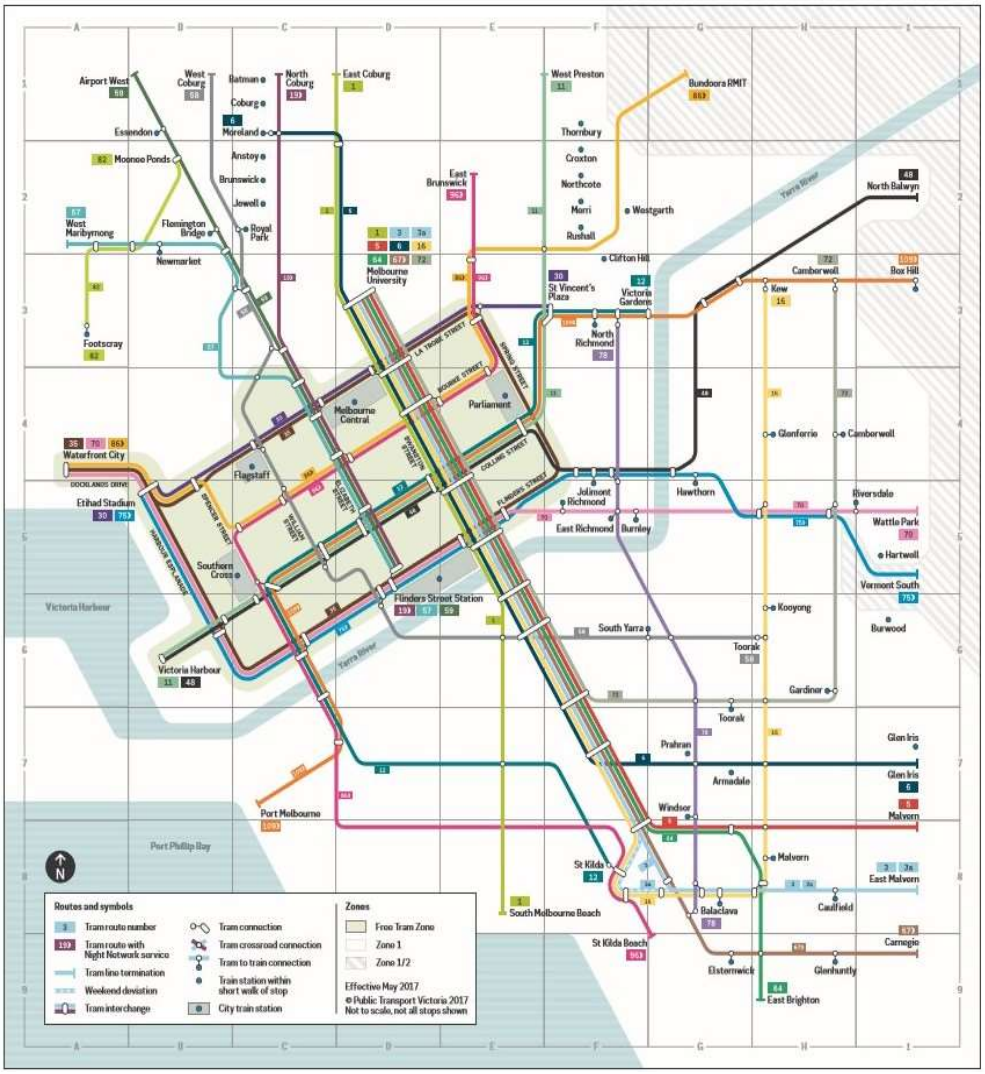

Figure 3.1: Map of Melbourne tram network (PTV 2018)

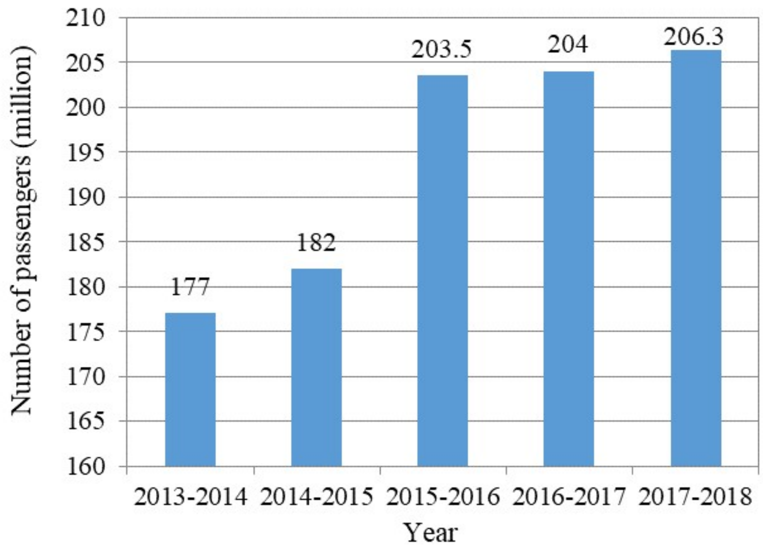

Figure 3.2: Melbourne tram patronage between 2013 and 2018 (PTV 2018)

In this research, the collected data cover six sequential years (2010 to 2015). Note that in this research a chord length (i.e., the length of the measure for collecting track geometry parameters) of 10 m was applied and the data was collected twice a year.

In this study, in order to extract the geometry deviations (track geometry parameters) associated with several years, data segmentation technique was applied. Data segmentation is the process of converting track record data, into track segments to facilitate the process of data matching. In this research, align with the chord length, a length of 10 m was selected for development of a track segment. Each track segment contains a track record with a specific identification code (a combination of track code and details of measurement location) along with track geometry parameters, track structural parameters and date of measurement. For data matching, by utilising SQL Query and the identification code, the values of the track geometry and structural parameters for six consecutive years of each track segment were collated and a multi-year dataset was built. In this research, more than 34,000 track segments have been processed and analysed. Figure 3.3 represents the development process for dataset preparation.

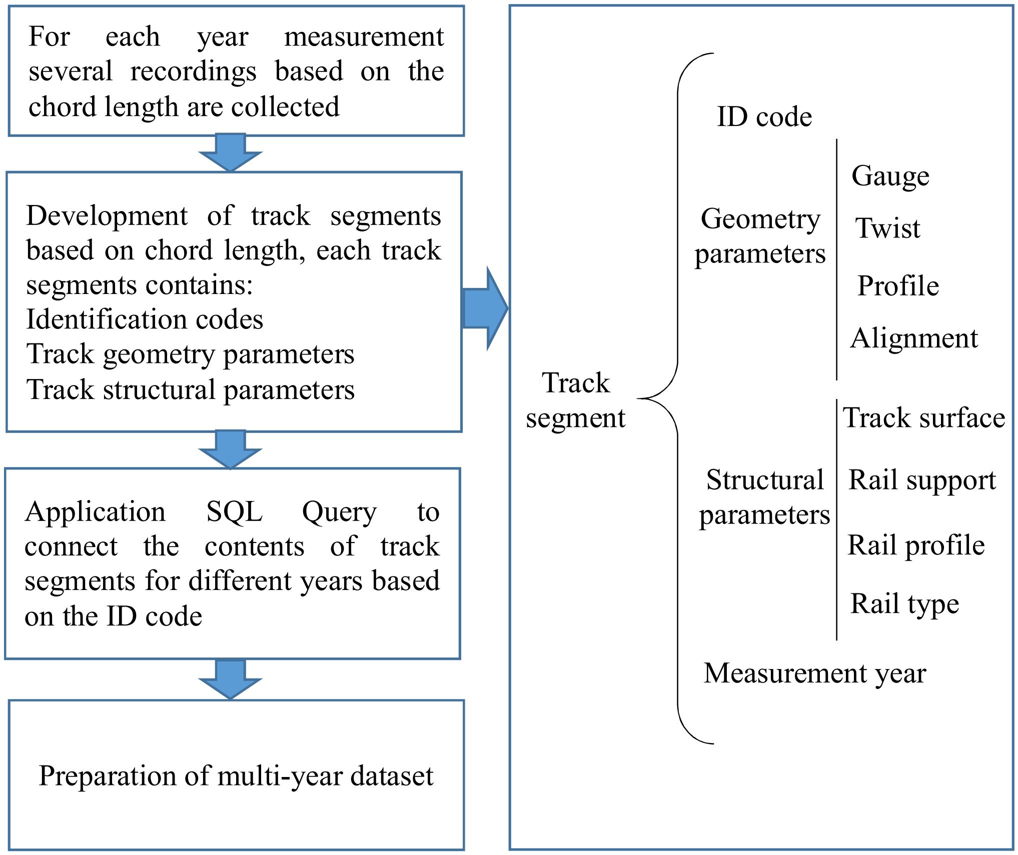

Figure 3.3: The process of dataset preparation

3.3. Research Framework

The research framework demonstrates a structure that is required to connect different part of the research including the gap of knowledge, research objectives, and resources in order to achieve the research aim. The research framework demonstrates the steps that should be taken. On other words research framework determines different stages of the research and their order from the starting point to the ending point. The research framework of this research consists of different steps which have been depicted in the following figure. It must be noted that the future chapters of the research have been categorised based on the steps provided in the research framework.

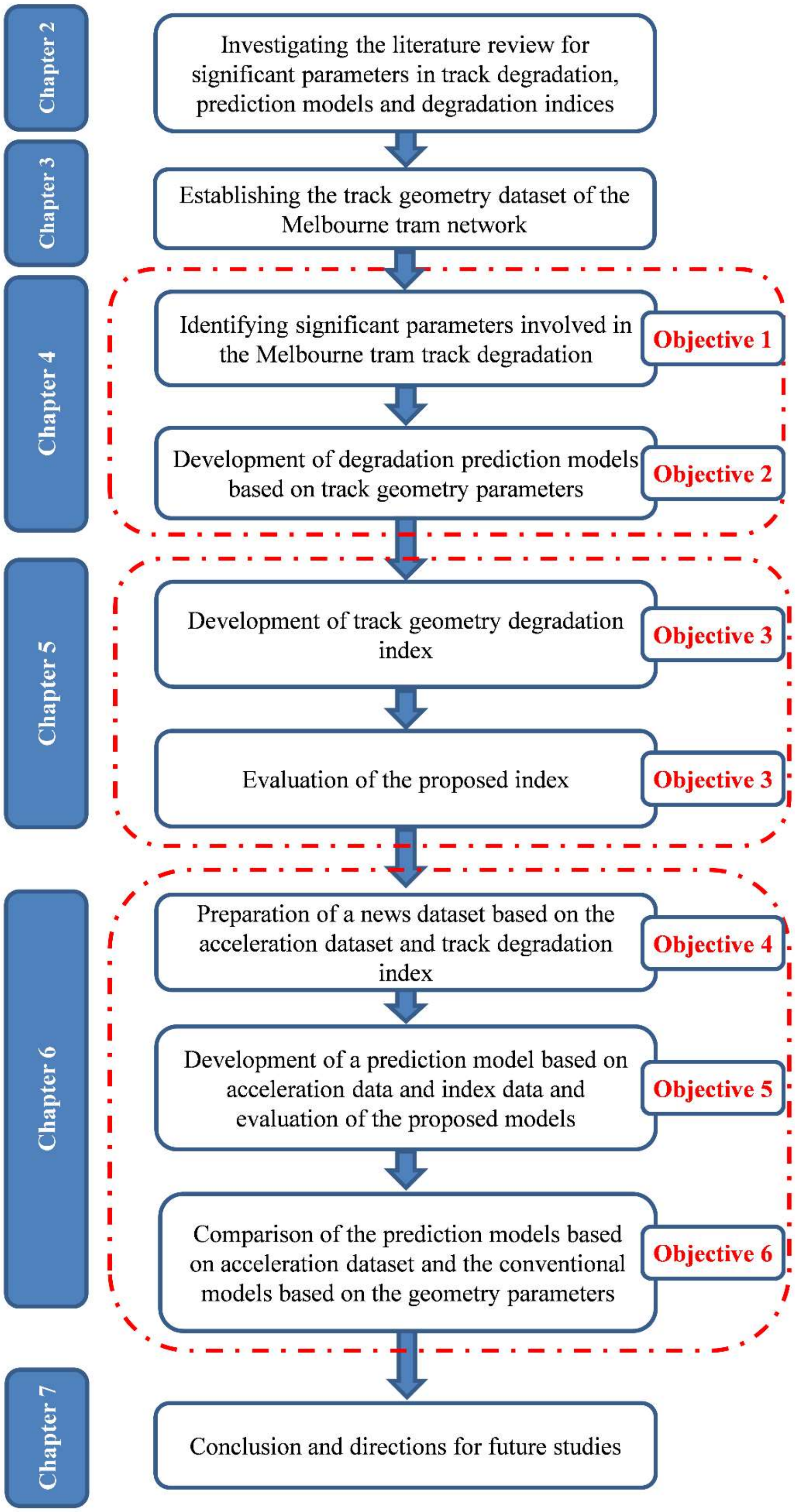

Figure 3.4: Research framework of the study

As illustrated in the flowchart for the research framework, after finalising the dataset of the research different steps needs to be completed.

In the first step, the importance of track geometry parameters and their effectiveness in the rail track geometry degradation process are examined. Without identifying the important factors in rail track degradation, development of prediction models to predict the future degradation of tram tracks is not feasible.

The second step is to develop degradation prediction models based on the influencing tram track parameters. Afterwards, the evaluation of the proposed models is undertaken. The outcomes of the model development can be considered as a benchmark for the rest of the study as they demonstrate the behaviour and function of the variables in the degradation process.

In the third step, the development of track degradation index as one of the main objectives of this study is targeted. The index represents the condition of track segments based on the historical geometry data and the condition of the degradation parameters. The proposed index needs to reflect the condition of influencing track geometry parameters together.

In the fourth step, the assessment of the proposed track degradation index is provided which aims to compare the effectiveness and accuracy of the index proposed in this research and the existing indices by applying them on the Melbourne dataset. For this purpose, several calculations along with statistical analyses need to be carried out.

In the fifth step, a new dataset based on the combination of the dataset derived from the track degradation index and dataset for tram acceleration data is prepared. Tram acceleration data as alternative data are captured from the movement of tram vehicles on rail tracks. The dataset created in this step consists of the value of track degradation index and its corresponding acceleration data.

In the six step of the framework, development of different models to predict the value of track degradation index based on the acceleration data is targeted. For this purpose, the dataset prepared in the previous step instead of the primary track geometry dataset is used. This step represents the aim of this study which is providing a cost-effective method to predict the rate of degradation in tram tracks.

In the seventh step of the framework, two methods described in this research for predicting tram track degradation is compared. In the first method, prediction of tram track degradation which is represented by the degradation index is relied on track geometry parameters and track structural parameters. The second method or the latest method is predicting the degradation index based on the acceleration data.

Finally, in the last step, the findings and the outcome of the research are discussed. In this regard, the role of different track parameters in tram track degradation and the evaluation of the proposed models are summarised. Moreover, innovative contributions to the research in different areas are presented. At the end of this section direction for future studies are provided. In future studies, the areas of the research that can be expanded are addressed.

3.4. Summary

In this chapter, first brief information related to the cases study of this research which is Melbourne tram network was provided including the number of trams, tram routes and tram stops. Also, information related to the Melbourne tram network patronage were provided such as the yearly patronage of the tram network. Afterwards, a dataset which is related to the case study was introduced. In this section track geometry parameters and track structural parameters involved in the Melbourne tram network were explained. Then the data segmentation technique which was used to integrate and link the dataset of various years was explained. The result of the data segmentation process in this research is a multi-year dataset based on six consecutive years. This dataset then will be used in the next chapter to identify effective parameters in tram track degradation prediction and also the degradation prediction modelling.

In the third section of this chapter, the research framework was explained. The research framework was designed based on the research aim, research objectives and the gap of the knowledge. In the research framework, the procedure required to reach the research aim was mapped out step by step in six consecutive steps. The research framework of the research specified the relationship between the objectives of the research and the important actions that must be accomplished to reach the aim of the research. These actions include statistical tests, different prediction model development, evaluation analyses, comparison of the results and the research conclusion.

CHAPTER 4

TRACK DEGRADATION PREDICTION MODEL DEVELOPMENT

4.1. Introduction

In this chapter, first the process for dataset preparation including data collection, data segmentation and data matching is presented. Afterwards, the significant parameters which have contribution to the degradation process of rail track are investigated. For this purpose, statistical analysis tests are applied to identify the significant parameters. After finalising the dataset, development of track degradation prediction models based on the above significant parameters are presented. In this regards, different models which have been investigated in the literature review section are examined and the models are developed. Based on evaluating measures, the performance of the proposed models is analysed. At the end of this chapter, the summary of the findings will be presented.

4.2. Identifying Effective Parameters

In this section effective parameters in tram track degradation prediction are identified then by applying a data filtering technique, outliers are discovered and removed.

Basically, for the establishment of a degradation prediction model, the existence of a meaningful correlation between previous and existing values of track geometry parameters is required. Involving parameters which are not statistically significant, can reduce the accuracy of the proposed model. In this study, the Pearson Correlation test as a measure of the relationship strength between two numeric variables has been used for both existing and previous values of the geometry parameters. Pearson correlation coefficients were calculated between -1 and 1. Coefficients closer to 1 and -1 have a greater correlation with the dependent variable and 0 means no correlation. Track geometry parameters which are correlated acceptably can be involved in the development of degradation models. Otherwise, they are not statistically significant and should be removed from the degradation model development. As demonstrated in Table 4.1 and according to the Pearson correlation coefficients and p-values, significant correlations between previous and existing deviations of gauge and twist parameters have been identified. Although the p-values of the test are less than 0.05 for the rest of the parameters (which demonstrates that the correlation between previous and existing values is statistically significant), their relationship is not strong enough for the development of the degradation models. Based on the result of this test, alignment, profile and cross-level parameters are removed from the track geometry dataset of the study.

Table 4.1: The results of the Pearson Correlation analysis.

Geometric deviations	**Pearson correlation coefficient**	**p-value**
Gauge	0.87	0.00
Twist	0.80	0.00
Alignment	0.17	0.00
Profile	0.13	0.00
Cross-level	0.25	0.00

Before using the dataset for further processes, a data filtration technique based on the mean value and standard deviation has been done to remove errors and out of range data in order to increase the accuracy of the dataset. A sample of changes in track gauge deviation of track segments over travelled distance for a particular track section is illustrated in Figure 4.1.

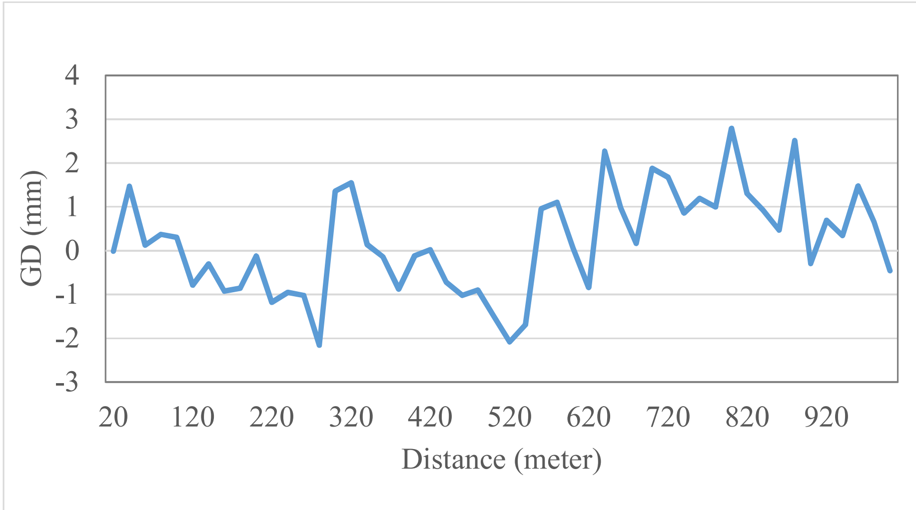

Figure 4.1: Changes in gauge deviation of track segments for a particular track section

For data filtration, determining the distribution patterns of the current dataset is useful. If the distribution of the dataset matches a normal distribution, 99.7% of the data will be within the distance of 3×SD from the mean value (DeGroot & Schervish 2012). For this purpose, the Shapiro-Wilk tests were conducted. This test is useful measures to check the possibility of a normal distribution.

After analysing different track sections, it was determined that the changes in GD and TD values mainly followed a normal distribution in which Shapiro-Wilk test: *p-value* > 0.05. The results of the test have been tabulated in Table 4.2.

Table 4.2: The results of Shapiro-Wilk test for GD and TD.

Geometric deviation	**p-value**
GD	0.819
TD	0.761

In Figure 4.2, the frequency histogram for a certain track section which follows a normal distribution is illustrated. In this section, track segments deviating from $\mu \pm 3 \times SD$ were identified and replaced with the closest values.

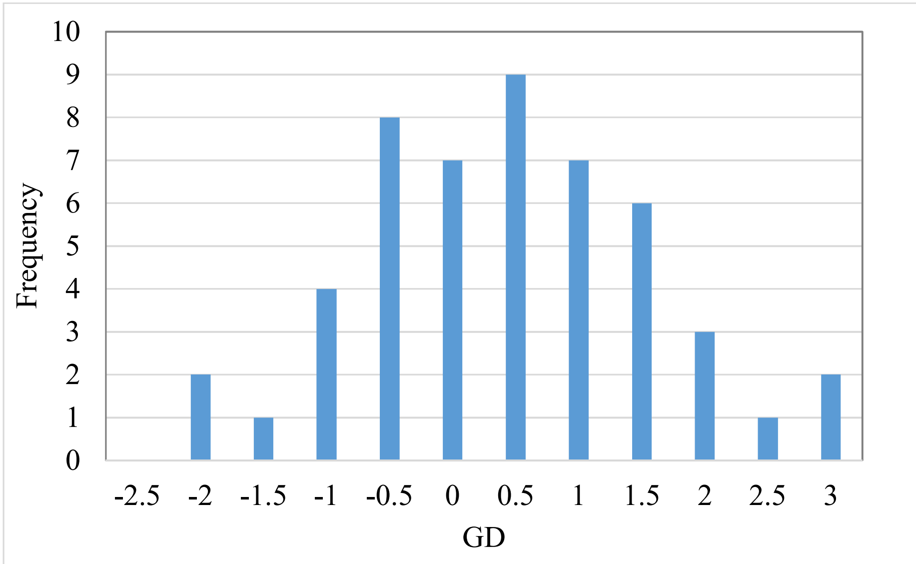

Figure 4.2: Frequency histogram of track segments for a certain track section

Before the model development, identifying the explanatory variables, which have a significant impact on the accuracy of the degradation models in predicting the dependent variable. Therefore, one-way ANOVA test (is a kind of ANOVA test that involves a single dependent variable and one independent variable with two or more categories) has been applied to the categorical variables and Pearson Correlation test has been applied to the continuous variables. Effective parameters contributing to the prediction of current gauge and twist deviation values (as dependent variables) have been investigated individually.

After applying the above tests, variables which have p-values less than the significance level (p-value<0.05) are statistically significant and probably worthwhile in the prediction of the dependent variable. These variables should be kept in the model development. Variables which have p-values greater than the significance level (p-value>0.05) are not statistically significant and should be removed from the model to increase the accuracy of the proposed models (Judd, McClelland & Ryan 2017).

Table 4.3 shows the results of the above statistical tests for gauge degradation prediction modelling and Table 4.4 shows the results of the above statistical tests for twist degradation prediction modelling:

Table 4.3: The results of statistical tests for the categorical and continuous variables for gauge deviation.

Test	Explanatory variables	Dependent variable	p-value
ANOVA	Rail type	Current gauge deviation	0.790
	Track surface		0.001
	Rail support		0.001
	Rail profile		0.060
Pearson Correlation	Previous gauge deviation		0.000
	MGT		0.620

Table 4.4: The results of statistical tests for the categorical and continuous variables for twist deviation.

Test	Explanatory variables	Dependent variable	p-value
ANOVA	Rail type	Current twist deviation	0.363
	Track surface		0.010
	Rail support		0.025
	Rail profile		0.070
Pearson Correlation	Previous twist deviation		0.000
	MGT		0.660

According to the results of the Tables 4.3 and 4.4, track surface and rail support as categorical variables (p-value< 0.05) are significant parameters in the prediction of the dependent variables which are gauge and twist deviations. Also, previous twist and gauge deviations as numerical variables are significant parameters in the prediction of the current twist and gauge deviations, respectively.

4.3. Model Development

In this section three different AI models including ANN models, SVM models and RFR models are developed and the evaluation processes are presented.

4.3.1. ANN Model Development

In this section, different structures of ANN models are developed to predict track geometry degradation based on gauge and twist parameters. The dependent variables which represent the level of degradation in the tram tracks are current deviation values of gauge and twist parameters. Explanatory variables are previous deviation values of twist and gauge parameters along with other structural parameters.

ANN is a branch of AI models inspired by biological neural networks. ANN consists of several independent interconnected neurons which can exchange message with each other through direct links called weighted connections. A neuron can process a result using values directly obtained from other neurons.

In this study, the Multi-layer Feed-forward network (MLF) is applied which has been widely used in engineering problems. As the MLF networks uses the multi-layers in its learning process, they can handle complex problems. Furthermore, MLF networks are very useful when the relationship between explanatory variables and the target variable is inherently non-linear. In this type of network, the neurons are organised in a layered architecture and the messages are transferred through the layers in a forward direction procedure (Fang et al. 2018; Paneiro et al. 2018). The link between the ith and jth neuron along with weight coefficient (ω_{ij}) and thresholds (v_i) are illustrated in Figure 4.3. The output of a neuron in a MLF network can be formulated as follows:

$$x_i = f\left(\sum_{j=1}^{n} \omega_{ij} \,.\, x_j + v_i\right) \tag{4.1}$$

Where x_i is the potential value of neuron ith, f represents the transfer function, ω_{ij} denotes the weight coefficient and v_i is the threshold coefficient.

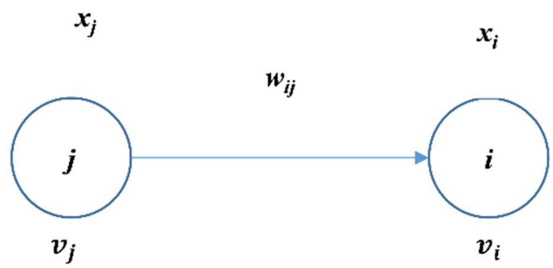

Figure 4.3: Link between the *i*th and *j*th neuron

In this study, the back propagation algorithm is used for training the dataset. In this algorithm, the error messages calculated from the difference between the actual and expected results are distributed back from the layer by layer. The training begins with random values of the weights. This algorithm updates and adjusts the weights of connections and biases in each iteration. The iterative process continues until the weights of connections are gradually optimised. In this study, two types of transfer functions including Logistic and Tangent Hyperbolicus (Tanh) are applied and checked. Transfer function is mathematical functions which their main function is to model the system's output for each possible input (Kwon 2017; Wang, Cheng & Li 2018).

In this study, a k-fold cross-validation technique has been used to optimise data sampling. The k-fold cross-validation can decrease bias in the learning process. Also, the variance of the resulting estimate is lowered as the value of k increases (by averaging over k different partitions).

In this study, a four-layered network (an input layer, two hidden layers and an output layer) with the different topology of networks (different number of neurons in the hidden layers) are created. Figure 4.4 illustrates a sample four-layered network with one input and output variables.

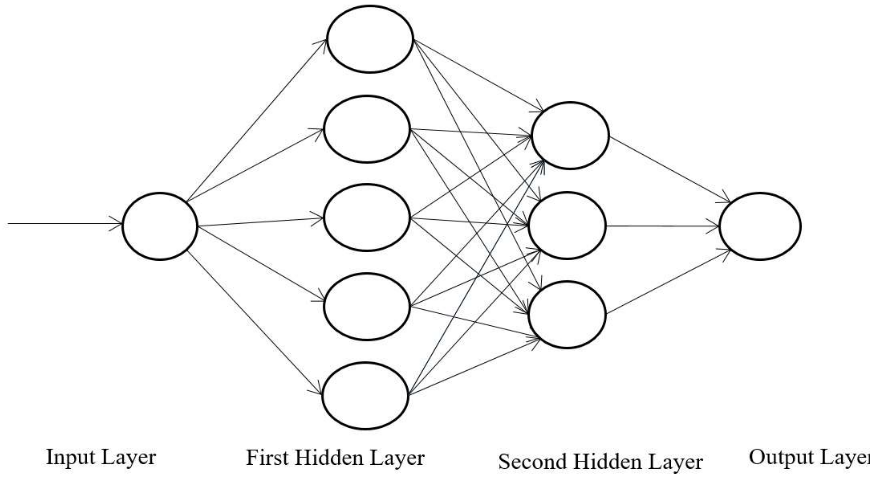

Figure 4.4: Topology of a four-layered network with one input variable

4.3.2. SVM Model Development

Support Vector Machine (SVM) as a branch of machine learning models, can be categorised into two major categories including Support Vector Regression (SVR) and Support Vector Classification (SVC). SVR applies the adaptive margin-based loss function and transfers the training data to higher dimensional feature space. The transform function is called kernel function. Kernel function transfers the data from non-linear space to linear space. Training support vector regression can be stated by $f(\mathrm{x}) = \omega . x + b$ regression function, where ω stands for adjustable weight parameter, x is the original input vector ($x \in R^n$) and b is the model parameter. The following optimisation problem must be solved to find the elements of the SVM:

$$\min \ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}\xi \qquad s.t. \quad \xi_i \geq 0 \tag{4.2}$$

Where ξ is a slack parameter and C represents the trade-off between the weight vector ($\|\omega\|$) and training error.

The optimisation problem can be solved by applying a Lagrangian multiplier method (Marković et al. 2015; Jiang et al. 2018).

In this study, three different types of kernel functions including Polynomial, Radial Basis Function (RBF) and Sigmoid are used and tested. RBF kernel function is defined as:

$$k(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad (4.3)$$

Where, γ controls the width of radial basis function and x_i represents the transformed input data.

In this study by applying Genetic Algorithm (GA), optimising SVM hyper-parameters (C and γ) are targeted. In this research, GA is selected for optimising the proposed models as it has robust global search capability. GA is applied in engineering problems to optimise solutions by imitating genetic evaluation process and using bio-inspired operators including selection, mating, crossover and mutation. Figure 4.5 illustrates the procedure of GA in a flowchart.

The process starts randomly with a set of individuals which is called a population. In the next step, the fitness function governs how fit an individual is. It assigns a fitness score for each individual. In the evaluation step, the fitness score will be used to determine whether needs to be reproduced or not. In the selection step, a portion of the existing population will be selected to breed a new generation. In this study, SVM hyper-parameters are the variables and Mean Squared Error (MSE) is the target of the fitness function.

The next step is the crossover which also called recombination. In this step, a genetic operator will be used to combine two individuals based on their genetic information to generate new offspring. In other words, it is a way to generate new solution stochastically from the existing population. The mutation is used to maintain diversity in the genetic population. As a result, one or more gene values will be altered in a chromosome from its initial state. In this step, the solution may change completely from the earlier solution. The algorithm ends if the population does not produce offspring which are meaningfully distinct from the previous generation. At this point, optimal results are provided. Otherwise, the main operation including selection, crossover and mutation will be repeated (Li et al. 2017; Falamarzi et al. 2019). In the proposed GA-SVM model, a k-fold cross-validation technique has been used to optimise data sampling.

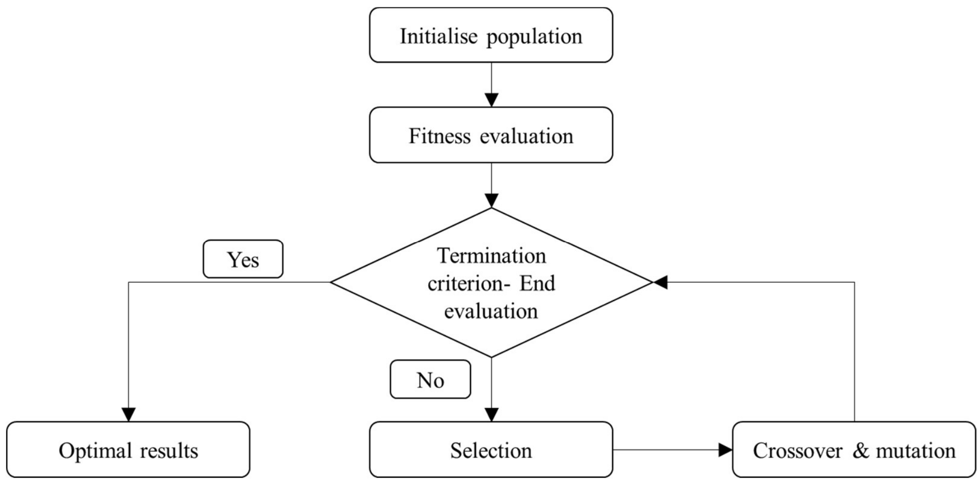

Figure 4.5: The process of GA

4.3.3. Random Forest Model Development

Random Forest (RF) model as a branch of Machine Learning (ML) is an ensemble-learning algorithm for both classification and regression problems. RF models are constructed based on the Decision Tree (DT) concept but with two distinguishing features including bagging and Random Subspace Method (RSM). Bagging (Bootstrap Aggregation) is a ML ensemble algorithm developed to improve the stability and accuracy of machine learning models by generating various bootstrap samples to avoid an overfitting problem. RSM algorithm reduces over-focusing of decision trees on features that appear to be highly descriptive or predictive in the training dataset. In this algorithm, a random sample of estimators with replacement (than the entire estimators) has been used to train decision trees in each bootstrap and at each node (Hua, Shen & Zhong 2017; Liu et al. 2018; Falamarzi et al. 2018b).

According to the literature, rapid data processing and high accuracy of RF method on large data are the advantages of RF models and the main reasons for selecting this method in this research. The RF development can be summarised in three main stages. In the first stage, based on the bagging method, several numbers of bootstrap samples extracted from the training dataset are randomly created with replacement. On average, each bootstrap consists of 63% of the training dataset. At each bootstrap sample, a decision tree is formed. In the second step, by considering RSM method each node of decision trees is split into sub-nodes. The number of estimators, which are randomly chosen at each node, is called *mtry*. In the last step, the majority or the average of outputs (from the bootstraps) represent the outcome of the prediction.

For optimising the model, calculating the Out of Bags (OOB) data is required. In this regard, at each time of bootstrap, the samples which are not picked to create a decision tree and involved in the learning process saved as OOB data. Then the OOB data is inputted into the developed decision trees and the prediction errors are calculated.

Then OOB prediction error should be aggregated. By changing the number of bootstrap samples (*ntree*) and *mtry*, the OOB error can be optimised. The model with the least OOB error is more reliable (Altman & Krzywinski 2017; Sharma et al. 2018). The flowchart in Figure 4.6 illustrates the process of the RF development based on a combination of bagging and RSM methods.

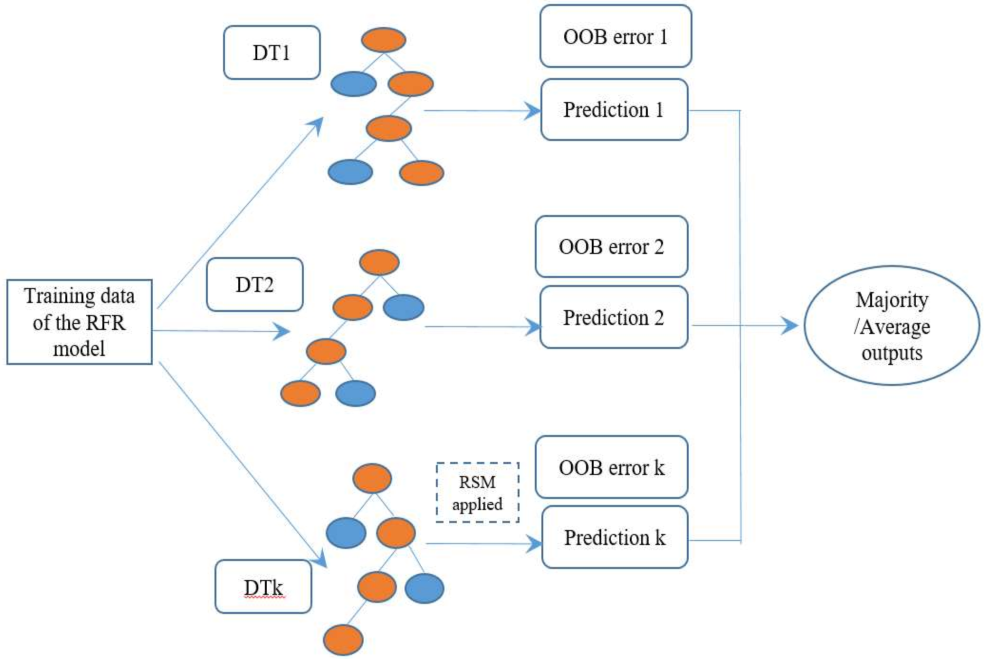

Figure 4.6: RF prediction process.

4.4. Results and Discussion

In this section, the results derived from the developed models are provided. For evaluation, different techniques can be applied to numerically assess the performance of the proposed models. As in this study, the dependent variable which is a continuous variable, adjusted R^2, which demonstrates the goodness of fit between the observed data and the predicted data (Equation 4.4), the Root Mean-Squared Error (RMSE) presented by Equation 4.5 and the Mean Absolute Percentage Error (MAPE) presented by Equation 4.6 have been used to assess the outcomes of the proposed models as follows (Were et al. 2015):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.4}$$

Where, R^2 represents the coefficient of determination, n is the number of validation samples, f_i denotes the value predicted by the model, y_i represents the observed data and $\bar{y}$ stands for the mean value of y_i.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \tag{4.5}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - f_i|}{y_i} \times 100 \tag{4.6}$$

For the assessment of the proposed models, greater values of adjusted R^2 (typically greater than 0.5) and lower values of RMSE and MAPE are desirable. It must be noted that as the tram track geometry dataset contains zero values, MAPE as a performance indicator is not applicable to assess the proposed models of this chapter.

4.4.1. ANN Results

In this section, the results of ANN model applied to both gauge and twist deviation, as well as structural parameters, are provided. For the model development, 75% of the dataset were dedicated to training the model and the rest is assigned for data validation and assessing the outcomes.

4.4.1.1. Gauge Deviation Results

Different models with different variable combinations, number of neurons and transfer function have been examined. Figure 4.7 shows the correlation between the real gauge deviation data and the data predicted by the proposed ANN models. The results of the most accurate model in terms of adjusted R^2 and RMSE have been tabulated in the following table.

Table 4.5: The results of the ANN model for gauge deviation.

Explanatory variables	Dependent variable	Neurons in hidden layers	Transfer function	Adjusted R^2	RMSE
Previous gauge deviation Track surface Rail support	Current gauge deviation	4,3	Tanh	0.79	1.67
		5,3	Tanh	0.79	1.67
		6,3	Tanh	0.79	1.67
		7,3	Tanh	0.80	1.65
		4,3	Logistic	0.79	1.66
		5,3	Logistic	0.79	1.67
		6,3	Logistic	0.79	1.67
		7,3	Logistic	0.80	1.65

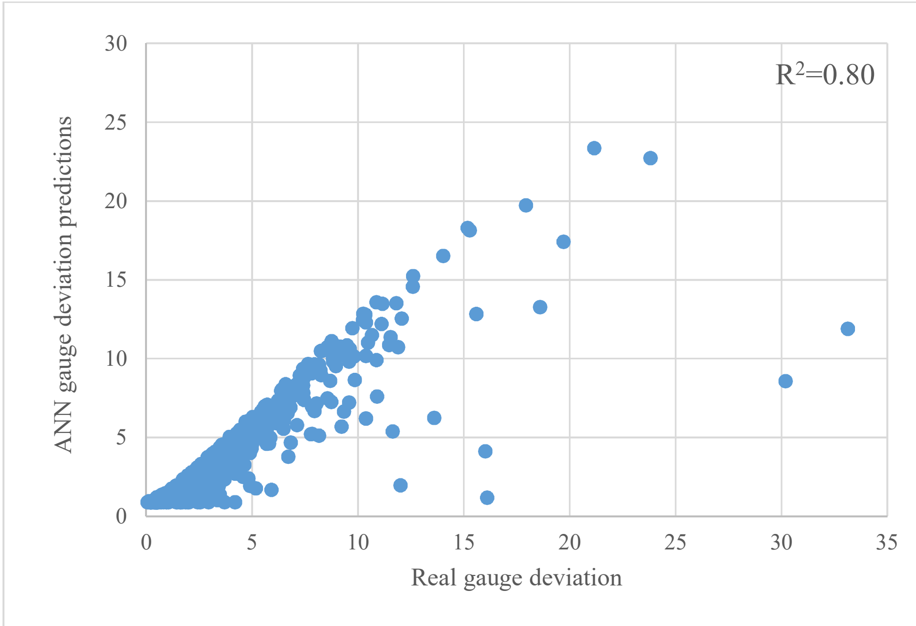

Figure 4.7: Real gauge deviation values against the ANN predictions.

As presented in Table 4.5 different ANN models with different specifications including the number of neuron in hidden layers and also transfer function have been developed and compared. The proposed models provided similar results in terms of the RMSE value as well as the coefficient of determination. The value of adjusted R^2 ranges between 0.79 and 0.80 and the value of RMSE ranges between 1.65 and 1.67 which demonstrates that the proposed models fit the data in the acceptable range and reasonable errors.

According to these results, the ANN models with neurons [7,3] in its hidden layers regardless of their transfer function have provided slightly better prediction compared to other developed models. According to the result of the ANN model development, an increase in the rate of previous gauge deviation and implementing steel sleeper and asphalt track surface can escalate the rate of degradation in tram tracks. Conversely, the degradation rate is mitigated by implementing concrete sleeper and concrete track surface. The weights of links for the proposed ANN model have been illustrated in Figure 4.8.

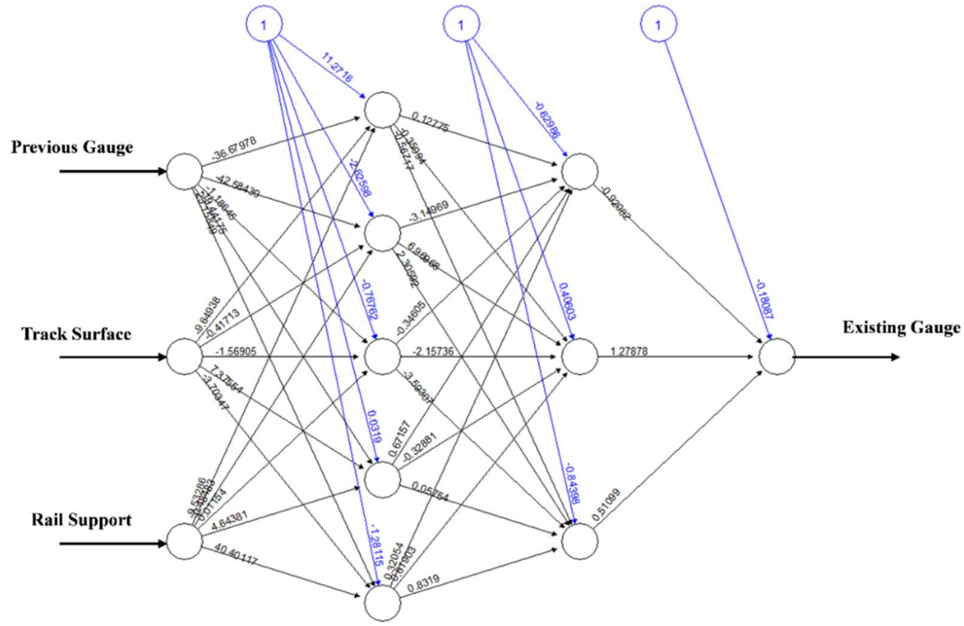

Figure 4.8: The weights of links for the proposed ANN model based on the gauge deviation.

4.4.1.2. Twist Deviation Results

In this section, different models with different variable combinations, the number of neurons and transfer function have been examined. Figure 4.9 shows the correlation between the real twist deviation data and the data predicted by the proposed ANN models. The results of the most accurate model in terms of adjusted R^2 and RMSE have been tabulated in Table 4.6 below.

Table 4.6: The results of the ANN model for twist deviation.

Explanatory variables	Dependent variable	Neurons in hidden layers	Transfer function	Adjusted R^2	RMSE
Previous twist deviation,track surface, rail support	Current twist deviation	4,3	Tanh	0.66	0.89
		5,3	Tanh	0.66	0.89
		6,3	Tanh	0.66	0.89
		7,3	Tanh	0.67	0.89
		4,3	Logistic	0.66	0.89
		5,3	Logistic	0.66	0.89
		6,3	Logistic	0.66	0.89
		7,3	Logistic	0.67	0.89

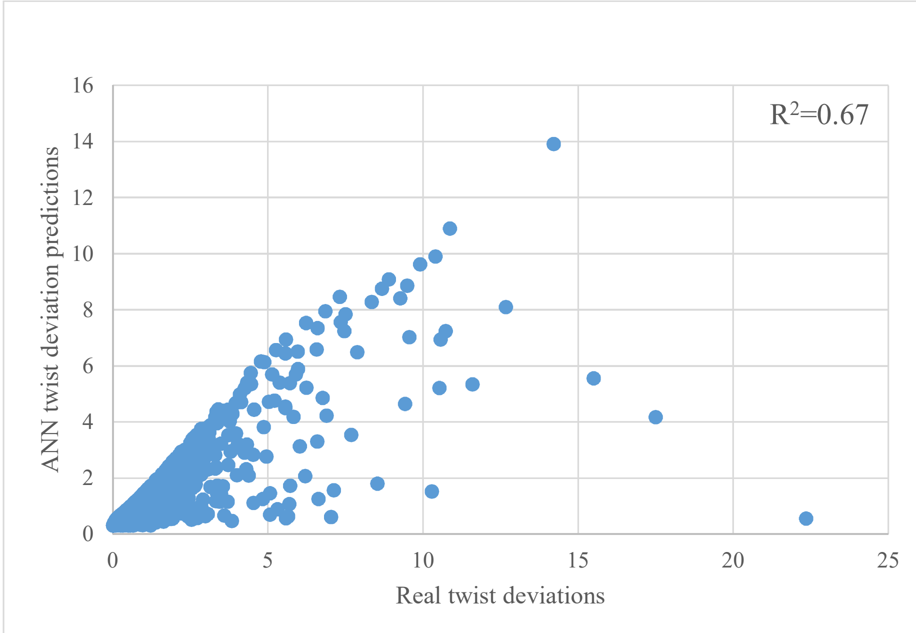

Figure 4.9: Real twist deviation values against the ANN predictions

As presented in Table 4.6 and similar to the outcome of gauge deviation prediction, the ANN models with [7,3] neurons in its hidden layers regardless of their transfer function have provided slightly better prediction compared to other developed models. According to the result of the ANN model development, an increase in the rate of previous twist deviation and implementing steel sleeper and asphalt track surface can escalate the rate of degradation in tram tracks. Conversely, the degradation rate is mitigated by implementing concrete sleeper and concrete track surface. The weights of links for the proposed ANN model have been illustrated in Figure 4.10.

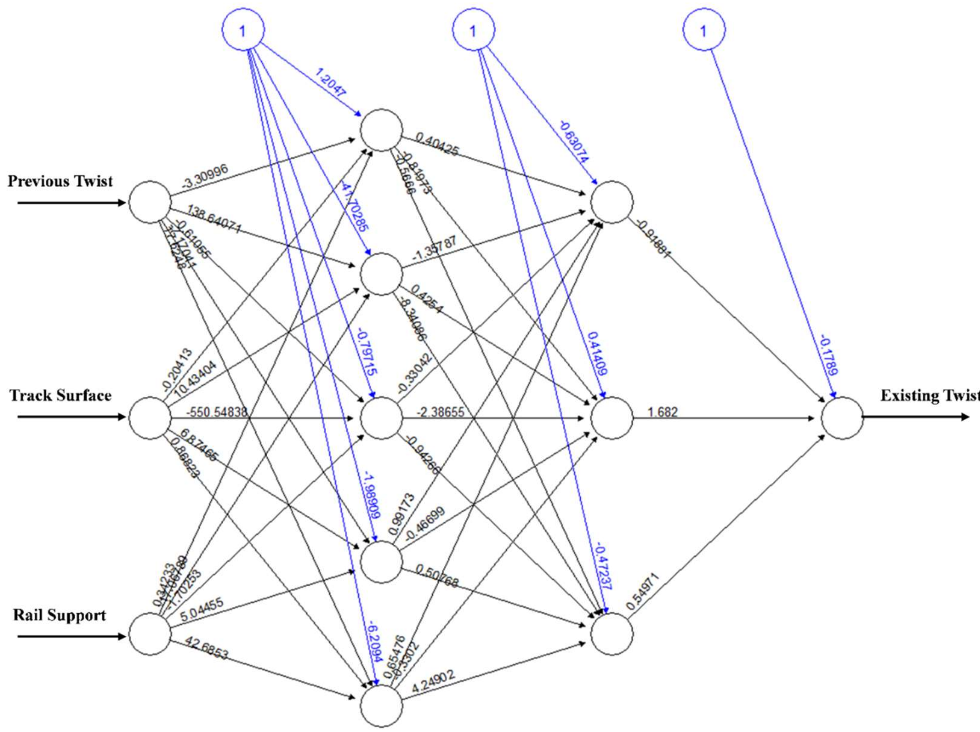

Figure 4.10: The weights of links for the proposed ANN model based on the twist deviation.

4.4.2. SVM Results

In this section, the results of the SVM model applied to both gauge and twist deviation, as well as structural parameters, are provided. For the model development, 75% of the dataset were dedicated to training the model and the rest is assigned for data validation and assessing the outcomes.

4.4.2.1. Gauge Deviation Results

In this section, similar to the ANN model, different models with different variable combination have been developed. Figure 4.11 shows the correlation between the real gauge deviation data and the data predicted by the proposed SVM models. The results for the model with the highest performance has been demonstrated in Table 4.7 below. As mentioned in Section 4.3.2, the hyper-parameters of the SVM model have been optimised with the GA technique to produce better results.

Table 4.7: The results of SVM model for gauge deviations.

Explanatory variables	**Dependent variable**	**Model type**	**Hyper-parameters**	**Kernel function**	**Adjusted R^2**	**RMSE**
Previous gauge deviation Track surface Rail support	Current gauge deviation	GA-SVM	*C*= 5.62 γ = 1.17	RBF	0.87	1.35
		SVM	*C*=1 γ =1		0.86	1.36

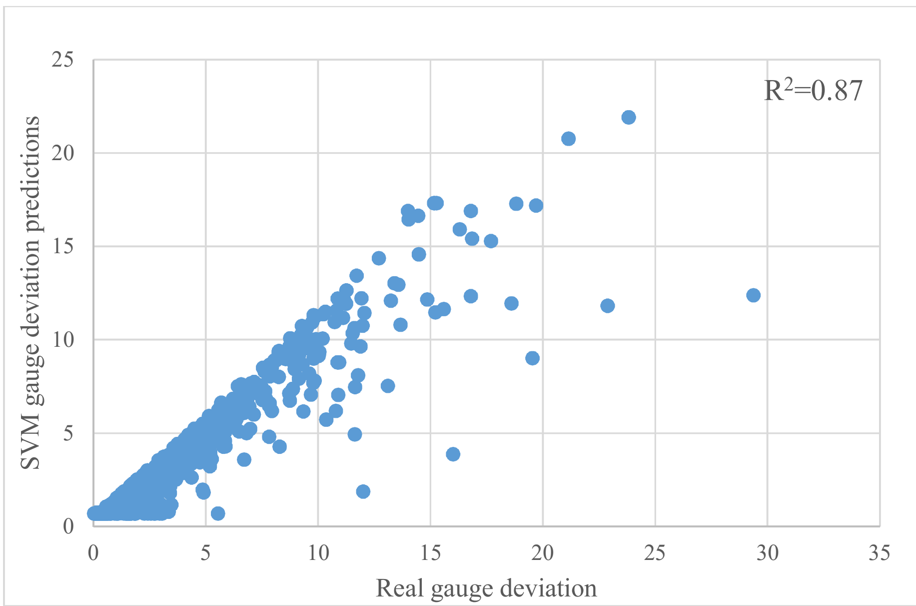

Figure 4.11: Real gauge deviation values against the SVM predictions.

As presented in Table 4.6 two different SVM models including SVM and GA-SVM models have been developed and the results are provided. Hyper-parameters in GA-SVM model has been optimised and the values were determined. RBF kernel function as a kernel function has been used in both proposed models at it provide more accurate results. According to this table, GA-SVM model has provided a slightly better prediction compared to the SVM model. In this model, adjusted R^2 equals 0.87 and the value of RMSE is 1.36 which demonstrates that the proposed models perform in acceptable ranges and similar to the ANN models.

In addition the result of the SVM model development implies that, an increase in the rate of previous gauge deviation and implementing steel sleeper and asphalt track surface can escalate the rate of degradation in tram tracks. Contrarily, the degradation rate is decreased by implementing concrete sleeper and concrete track surface.

4.4.2.2. Twist Deviation Results

In this section, sample datasets for training and testing of the model have been created. Different models with different variable combination have been developed. Figure 4.12 shows the correlation between the real twist deviation data and the data predicted by the proposed SVM models. The results for the model with the highest performance has been demonstrated in Table 4.8 below. As mentioned in Section 4.3.2, the hyper-parameters of the SVM model have been optimised with the GA technique to produce better results.

Table 4.8: The results of SVM models for twist deviations.

Explanatory variables	**Dependent variable**	**Model type**	**Hyper-parameters**	**Kernel function**	**Adjusted R^2**	**RMSE**
Previous twist deviation Track surface Rail support	Current twist deviation	GA-SVM	C= 6.31 γ= 0.75	RBF	0.66	0.94
		SVM	C=1 γ=1		0.65	0.96

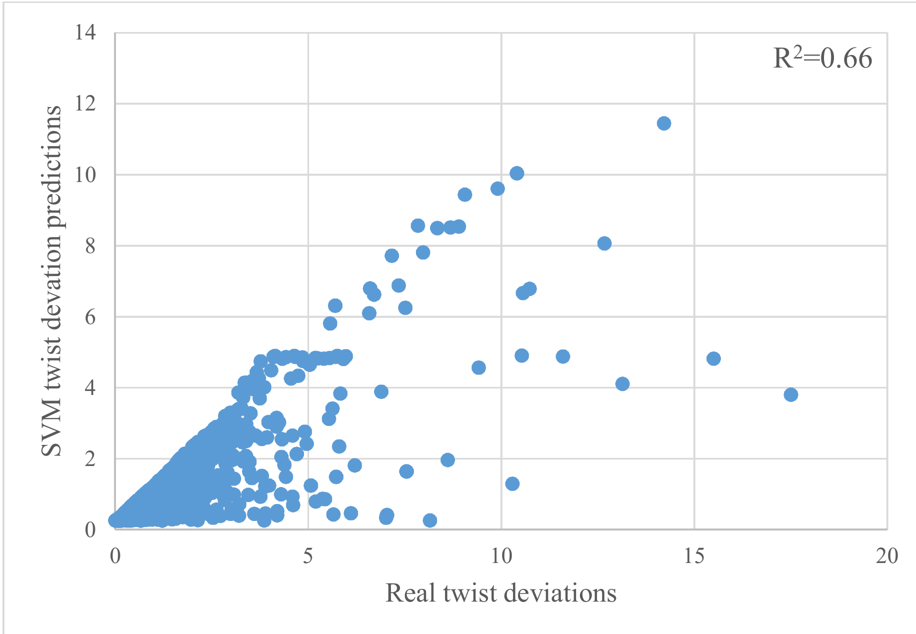

Figure 4.12: Real twist deviation values against the SVM predictions.

As presented in Table 4.7 and similar to gauge deviation predictions, GA-SVM model has provided better results in the prediction of future twist deviations. In this model, the RBF kernel function has been used as the transfer function. In this model, adjusted R^2 equals 0.66 and the value of RMSE is 0.94. Based on these results provided, an increase in the rate of previous twist deviation and implementing steel sleeper and asphalt track surface can increase the rate of degradation in tram tracks. Contrarily, the degradation rate is reversed by implementing concrete sleeper and concrete track surface.

4.4.3. RFR Results

In this section, the results of the RFR model applied to both gauge and twist deviation, as well as the structural parameters, are provided. Similar to ANN and SVM model developments, in this section, 75% of the dataset were dedicated to training the model and the rest is assigned for data validation and assessing the outcomes.

4.4.3.1. Gauge Deviation Results

In this section, the results of the RFR model development along with the evaluation of the model are provided. In this section, different models with different explanatory variables, *ntree* and *mtry* values have been developed. Figure 4.13 shows the correlation between the real gauge deviation data and the data predicted by the proposed RFR models. The results of the model evaluation with respect to adjusted R^2 and RMSE are demonstrated in Table 4.9 below.

Table 4.9: The results of the RFR models for gauge deviation

Explanatory variables	Dependent variable	*ntree*	*mtry*	Adjusted R^2	RMSE
Previous gauge deviation Track surface Rail support	Current gauge deviation	50	2	0.92	1.60
		100	2	0.93	1.57
		50	3	0.92	1.19
		100	3	0.93	1.16

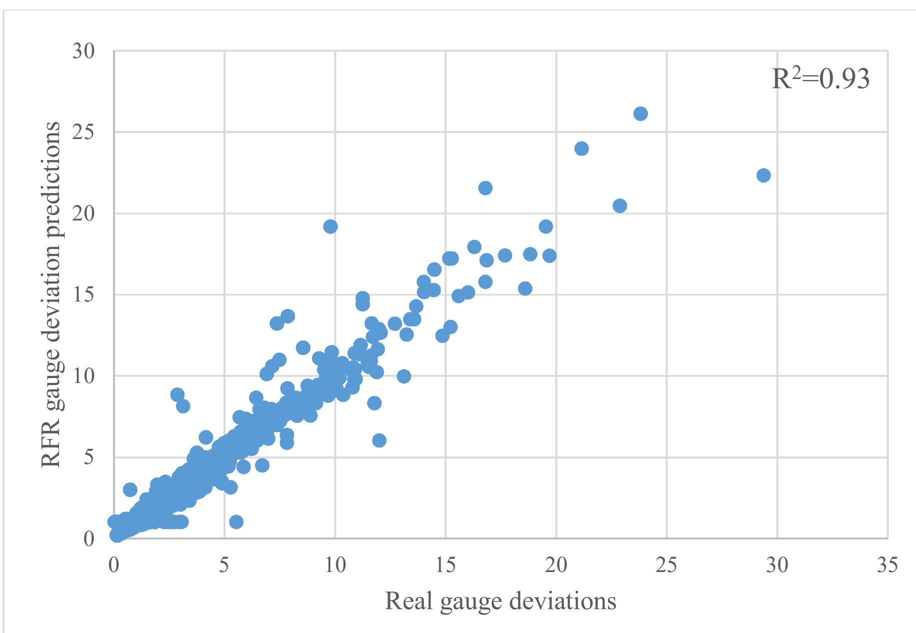

Figure 4.13: Real gauge deviation values against the RFR predictions.

As presented in the above table, different RFR models with regard to *ntree* and *mtry* have been developed and the results are provided. Two different values for *ntree* parameter and two different values for *mtry* paramater have been tested. According to the results, the model with *mtry* of 3 and *ntree* 100 has provided better prediction compared to other alternatives. As presented in this table, the proposed RFR model has adjusted R^2 of 0.93 and RMSE of 1.19.

In Figure 4.14, the OOB error rate changes versus the number of trees (*ntree*) and for *mtry* 3 has been shown. For the proposed models, the OOB error rate has been decreased gradually by increasing the number of trees. Then the OOB error rate stabilises around the number of 100 trees.

Based on the results of the analysis, there is a direct relationship between previous gauge deviation and current gauge deviation. The degradation rate of tram track will be reduced when concrete sleeper and concrete track surface have been used in the tram track infrastructure. These results are consistent with previous findings of rail track degradation which examined the degradation rate based on track infrastructure (Laryea et al. 2014; Falamarzi et al. 2017; Falamarzi et al. 2018a).

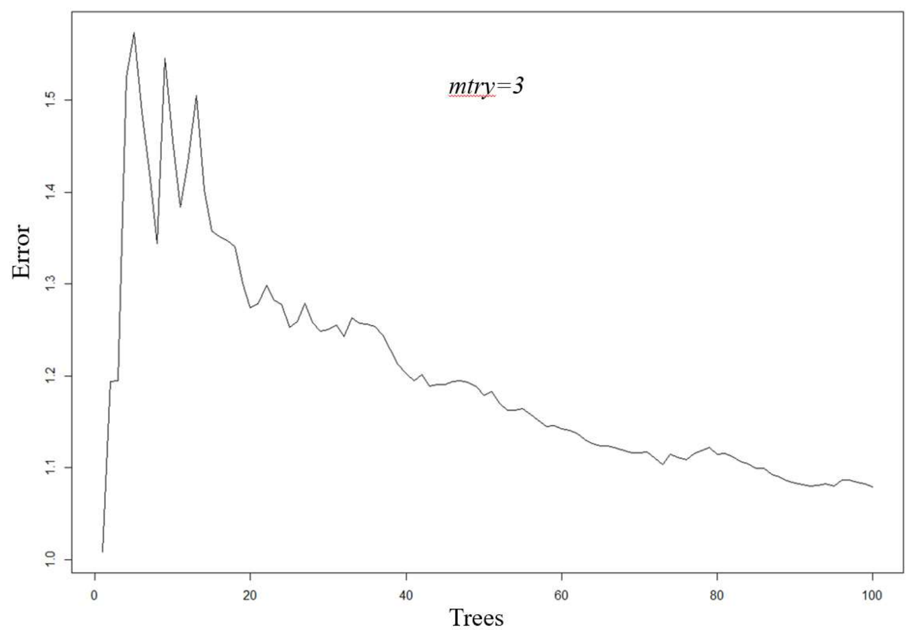

Figure 4.14. OOB error rate of the proposed RFR model.

4.4.3.2. Twist Deviation Results

In this section, the results of the RFR model development along with the evaluation of the model are provided. In this section, different models with different explanatory variables, *ntree* and *mtry* values have been developed. Figure 4.15 shows the correlation between the real twist deviation data and the data predicted by the proposed RFR models. The results of the model evaluation with respect to adjusted R^2 and RMSE are demonstrated in Table 4.10 below.

Table 4.10: The results of the RFR models for twist deviation.

Explanatory variables	Dependent variable	*ntree*	*mtry*	Adjusted R^2	RMSE
Previous twist deviation Track surface Rail support	Current twist deviation	50	2	0.82	0.84
		100	2	0.83	0.85
		50	3	0.83	0.78
		100	3	0.83	0.77

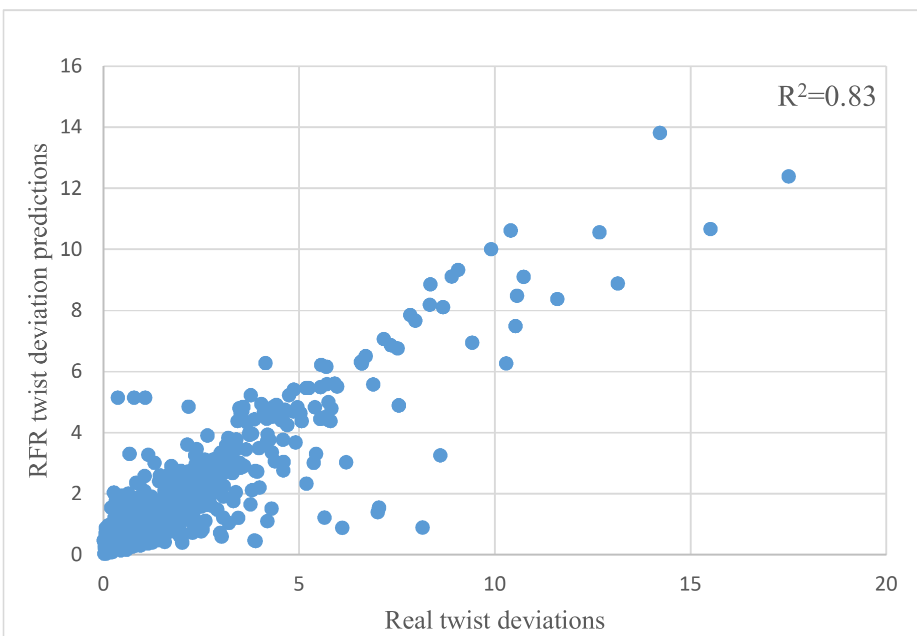

Figure 4.15: Real twist deviation values against the RFR predictions.

Similar to twist degradation prediction models with regard to *ntree* and *mtry* have been developed and the results are provided. Two different values for *ntree* parameter and two different values for *mtry* parmater have been tested. According to the results, the model with *mtry* of 3 and *ntree* 100 has provided better prediction compared to other alternatives. As presented in this table, the proposed RFR model has adjusted R^2 of 0.83 and RMSE of 0.77.

The OOB error rate changes versus the number of trees (*ntree*) and for *mtry* 3 has been shown in Figure 4.16. As illustrated in this figure, OOB error rate has been dropped gradually by increase in the number of trees. Then the OOB error rate stabilises around the number of 100 trees.

Based on the results of the analysis, a direct relationship between previous gauge deviation and current gauge deviation exists. The degradation rate of tram track will be decreased when concrete track surface and concrete sleeper have been installed in the tram track infrastructure.

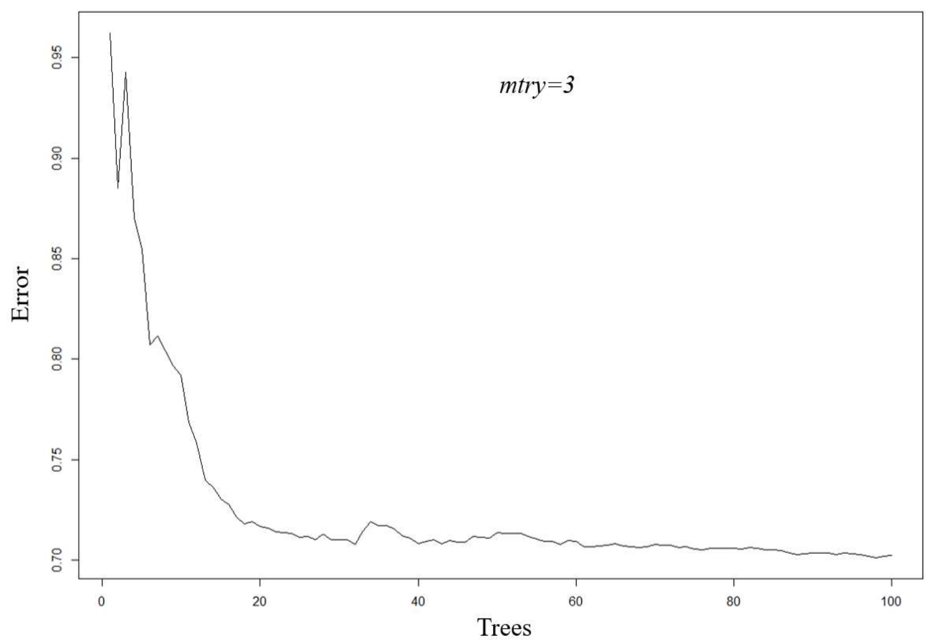

Figure 4.16: OOB error rate of the proposed RFR model

4.5. Summary

In this chapter, influencing parameters in tram track degradation prediction modelling as well as the Machine Learning (ML) approaches to predict future rail track degradation were explained and developed. By applying statistical tests influencing parameters in tram track degradation have been identified. In this regard, track surface and rail support as effective structural parameters and previous gauge and twist deviations as geometry parameters were engaged in tram track degradation modelling. Both future gauge and twist deviation have been predicted separately based on the above parameters. Artificial Neural Networks (ANN), Support Vector Machine (SVM) and Random Forest Regression (RFR) models are effective approaches which have been used successfully in different rail track degradation practices. In this research, a k-fold cross-validation technique has been used to optimise the parameters used in ANN and SVM models. RFR model is optimised version of decision tree approach by utilising bagging and RSM methods.

According to the results of the proposed models and based on the performance indicators including the coefficient of determination and RMSE, the outcome of the models are within reasonable and acceptable ranges. In this regard, RFR models provide better predictions in terms of the performance indicators compared to ANN and SVM models. The outcomes of this chapter can be considered as a benchmark for the next stages of the research framework as the gauge and twist parameters have successfully played an important role in providing a picture of tram track degradation.

CHAPTER 5

DEVELOPMENT OF TRAM TRACK DEGRADATION INDEX

5.1. Introduction

This chapter presents the procedure of developing a tram track degradation index. Track degradation indices can be used as an indicator of rail condition concerning the risk of damage or failure over a period of time. The index can be applied in establishing a sustainable tram track maintenance management system. This index can be also utilised later in rail degradation prediction modelling to represent the future condition of the rail track. Previous studies have mainly focused on conventional rail track degradation indices and light rail tracks have not been addressed properly. In this research dataset of the Melbourne tram network is used. In this research, based on the statistical analysis conducted in the chapter for data and model development, track geometry parameters which are statistically significant in the track degradation prediction models are used for the development of the proposed index. For the purpose of evaluation, the predictability performance of the index proposed in this research is compared with the three major existing indices in the literature. At the end of the chapter, the conclusions are provided.

5.2. Index Development

Degradation indices for rail track represent the quality of rail tracks with regard to track geometry parameters. The proposed index should be easy to use and at the same time reflect the geometry condition of track segments properly. Based on the findings obtained from the literature review, two main parameters including the mean value of geometry deviation in different measurements (years) and also the average differential geometry deviation have a strong contribution in the development track degradation indices. These parameters have been used and implemented separately in different degradation indices.

The mean value of the geometry deviations is an essential factor for formulating the degradation index, as larger values demonstrate more deviation from the original geometry parameter and consequently more risk of track failure compared to rail tracks with the lower mean values. Furthermore, the role of the average differential geometry deviation is important. It can be calculated by dividing the sum of the absolute value of differences between two consecutive geometry deviations by the total number of data collection years for a specific track segment. It is notable that two different track segments can have an accidentally equal mean value of geometry deviation, but the one with a larger differential geometry deviation reflects a faster rate of degradation in track geometry than the segment with a lower differential geometry deviation. On the other hand, two different track segments may have an equal average differential geometry deviation, but the one with a higher mean value represents a faster rate of degradation.

In this research, with regard to the importance of the above parameters, a new track degradation index has been proposed based on the combination of the mean value of the geometry deviation and the average differential geometry deviation. By utilising these parameters, the proposed index can benefit from the advantages of each of them. The proposed track degradation index is formulated as follows:

$$TDI_i = \mu_i + \lambda_i \tag{5.1}$$

Where, TDI_i represents the track degradation index based on the geometry deviation values including Gauge Deviation (GD), Twist Deviation (TD), Alignment Deviation (AD), Profile Deviation (PD) and Cross-level Deviation (CD) for the track segment *i*. μ_i is the mean value of the geometry deviations for the track segment i and λ_i represents the average differential geometry deviation for the track segment *i*.

The mean values of the geometry deviation of track segments for the consecutive years (μ_i) can be calculated using the following formula.

$$\mu_i = \frac{1}{m}\sum_{t=1}^{m} G_{dev_t} \tag{5.2}$$

Where, μ_i is the mean value of the geometry deviations for the track segment *i*, G_{dev_t} represents the geometry deviation of track segment i in year *t*, and m denotes the number of years for which data were collected.

The average differential geometry deviation (λ_i) was also included in the index formulation. λ_i can be determined by dividing the sum of the absolute value of differences between two consecutive geometry deviations by the total number of data collection years for the track segment *i* as follows:

$$\lambda_i = \frac{1}{m}\sum_{t=1}^{m-1} \left|G_{dev_{t+1}} - G_{dev_t}\right| \tag{5.3}$$

Where, λ_i represent the average differential geometry deviation of the segment *i*, and G_{dev_t} and $G_{dev_{t-1}}$ represent two consecutive geometry deviation values for the track segment *i*.

Once the values of TDI_i are calculated individually for all the track geometry parameters, then the Overall Track Degradation Index (OTDI) can be obtained. The value of OTDI can be calculated by using the following formula:

$$OTDI_i = \frac{\sqrt{a \times TDI_{GD}{}^2 + b \times TDI_{TD}{}^2 + c \times TDI_{AD}{}^2 + d \times TDI_{PD}{}^2 + e \times TDI_{CD}{}^2}}{a + b + c + d + e} \tag{5.4}$$

Where, $OTDI_i$ represents the overall track degradation index for the segment *i*, TDI_{GD}, TDI_{TD}, TDI_{AD}, TDI_{PD} and TDI_{CD} denotes TDI value based on the GD, TD, AD, PD and CD respectively for the track segment *i* and *a, b, c, d* and *e* are constant coefficients. The constant coefficients equal 1 while the geometry parameters associated with each of these coefficients play a role in the development of the index. Otherwise these coefficients will be zero. As discussed above, regarding the number of track geometry parameters involved in a certain dataset, Equation 5.4 can be modified.

In the next section, based on the dataset of the Melbourne tram network, the proposed index is implemented and assessed.

5.3. Case Study Development

For the index development, the dataset of the study which has been previously explained in Chapter 3 is used. In this study, the conditions of rail track were examined concerning GD (for both positive gauges, where rail heads diverge from the centreline of the track and negative gauge, where rail heads converge toward the centreline of the track) and TD. Consistent with the literature (Kopf et al. 2009; Wilson & Ker 2013; Guler 2014) and based on the GD and TD values, the rail conditions were classified into different levels and the values associated with TDI were calculated. Table 5.1 presents the track condition based on the TDI values for the gauge parameter and Table 5.2 presents the track condition based on the TDI values for the twist parameter. Finally, Table 5.3 presents the overall track condition based on the OTDI values.

Table 5.1: Track condition based on TDI_{GD}.

GD	limit	TDI
Positive GD	IAL	$35 \leq TDI$
	IL	$30 \leq TDI < 35$
	AL	$25 \leq TDI < 30$
Negative GD	IAL	$TDI \leq -11$
	IL	$-11 \leq TDI < -9$
	AL	$-9 \leq TDI < -7$

Table 5.2: Track condition based on TDI_{TD}.

limit	TDI
IAL	$24.5 \le TDI$
IL	$17.5 \le TDI < 24.5$
AL	$14 \le TDI < 17.5$

Table 5.3: Overall Track condition based on OTDI.

GD	limit	OTDI
Positive GD	IAL	$21.5 \le TDI$
	IL	17.5≤TDI<21.5
	AL	14.5≤TDI<17.5
Negative GD	IAL	TDI≤-13.5
	IL	-9.5≤TDI<-13.5
	AL	-7.5≤TDI<-9.5

In the above tables, Immediate Action Limit (IAL) denotes the value if exceeded, train speed restrictions or prompt correction of track geometry will be required. Intervention Limit (IL) denotes the value if exceeded, corrective maintenance operations are required to avoid IAL. AL or Alert Limit denotes the value if occurs, track geometry condition should be analysed and the planned maintenance operations are required regularly.

According to Table 5.3, for example, if OTDI values for positive GD are ranged between 14.5 and 17.5, planned maintenance operations should be scheduled. If OTDI values for positive GD are ranged between 17.5 and 21.5, corrective maintenance operations should be scheduled. Lastly, if OTDI values exceed 21.5, prompt correction of rail track is required.

In order to better understand how the proposed indices can be used to determine the condition of rail tracks regarding failure risk, a sample track section with the length of 100 m (containing 5 track segments) from the case study was examined. In this context μ_i and λ_i were calculated using Equation 5.2 and Equation 5.3, respectively. *TDI* (Equation 5.1) and *OTDI* (Equation 5.4) were then calculated and the results are tabulated in Table 5.4.

Table 5.4: TDI and OTDI values obtained for track segments of a sample track section

Seg. No.	Geo. Par.	Geometry deviation (G_{dev}) values measured over 6 years						μ_i	λ_i	*TDI*	*OTDI*
		2010	2011	2012	2013	2014	2015				
1	GD	5.61	1.76	1.69	4.97	7.38	4.95	4.39	2.01	6.40	5.32
	TD	6.45	7.10	7.67	7.65	8.38	10.05	7.88	0.61	8.49	
2	GD	12.07	10.72	11.92	5.69	3.2	10.21	8.97	3.05	12.02	6.34
	TD	3.76	3.84	3.88	3.63	4.05	4.24	3.90	0.16	4.06	
3	GD	4.81	4.76	3.85	5.65	2.92	3.97	4.33	1.09	5.42	3.32
	TD	3.16	3.78	3.10	3.11	3.27	4.17	3.43	0.40	3.83	
4	GD	3.49	2.77	2.07	3.59	3.07	4.23	3.20	0.77	3.97	3.47
	TD	4.64	4.69	4.80	5.45	5.69	6.79	5.34	0.36	5.70	
5	GD	8.75	7.32	5.51	7.59	8.88	9.95	8.00	1.28	9.28	5.17
	TD	3.20	3.49	4.19	4.22	5.09	5.18	4.23	0.33	4.56	
6	GD	2.60	3.45	4.18	3.90	5.52	5.80	4.24	0.63	4.87	4.16
	TD	5.55	4.87	6.30	6.10	7.12	7.15	6.18	0.56	6.74	
7	GD	3.12	4.05	5.60	4.90	4.80	5.12	4.60	0.60	5.20	4.90
	TD	7.30	8.12	8.39	7.52	7.99	8.05	7.90	0.42	8.31	
8	GD	7.12	7.15	8.06	8.38	9.12	9.57	8.23	0.41	8.64	5.11
	TD	3.33	3.45	5.36	5.81	5.90	6.12	5.00	0.47	5.46	
9	GD	4.44	4.78	5.63	6.12	6.51	5.80	5.55	0.46	6.01	3.48
	TD	2.78	2.90	3.20	2.77	3.65	3.89	3.20	0.33	3.53	
10	GD	3.18	3.74	5.26	5.88	5.60	6.01	4.95	0.57	5.51	4.63
	TD	5.26	5.49	6.07	7.79	8.01	8.62	6.87	0.56	7.43	

According to Table 5.4, the *TDI* values calculated from the GDs and TDs range between 3.53 and 12.02. OTDI values range between 3.32 and 6.34. According to

Table 5.3 and based on the calculated OTDI values these segments are categorised in no safety concern zone.

For instance, Segment 2 has the highest O*TDI,* which implies a higher risk of rail track failure in the long term compared with other segments and this segment should, therefore, be treated with higher priority than the others. Segments 1 and 3 have accidentally similar mean values of gauge deviation, but as the average differential gauge deviation of Segment 1 is higher, its TDI_{GD} value is larger. On the other hand, for example, Segments 4 and 5 have almost similar average differential twist deviation, but as the mean value of twist deviation of Segment 4 is higher, consequently its TDI_{TD} value is larger. These examples demonstrate the importance of the average differential geometry deviation as well as the mean value of geometry deviation in the proposed degradation index.

5.4. Index Evaluation

In this section, the evaluation of the proposed index implemented on the Melbourne tram network is presented. As discussed, the main roles of degradation indices are to represent the current condition of rail track geometry parameters of track segments as well as their future conditions. The future value of a degradation index can be used by rail organisers and operators to address preventive maintenance strategies prior to rail track failures. One of the potential ways to obtain the future value of a degradation index is to predict it by applying existing data. For this purpose, the predictability performance of the proposed index should be analysed. An index with greater predictability performance can be used by predictive models to provide predictions more effectively. In this regard, the correlation between the consecutive values of a track degradation index for a specific route or network is essential (e.g. $OTDI_{2015}$ & $OTDI_{2014}$). A stronger correlation between the values of an index in consecutive years demonstrates that the index has greater predictability performance.

In this research for carrying out the evaluation, the correlation between the current value of OTDI index and the previous value of OTDI is examined. The current degradation index value is obtained based on the geometry deviations measured up to the current year. While for the previous degradation index value, geometry deviations are processed up to the last year. In order to compare the performance of the proposed index and other studied indices, three major degradation indices including the index based on the standard deviation, the index based on the average of squared differential geometry deviation (Amtrak) and the index based on the mean value and standard deviation are included. The structure of the above indices has been explained comprehensively in Chapter 2 (Review of the existing studies).

To evaluate the performance of the indices, the Melbourne tram network dataset is used. In this dataset, the values of the previous and current OTDI, as well as other three indices for each track segments, are calculated. The Pearson correlation analysis was applied. Figure 5.1 to 5.4 illustrate the correlation between the current and previous values of the indices. Tables 5.5 presents the results of the analysis in terms of the performance indicators including Pearson correlation coefficient and Root Mean Squared Error (RMSE).

As illustrated in the following figures, indices, where data points are more scattered and located with more distances, have lower correlation coefficients than indices with dense population. According to Table 5.5 and based on Pearson correlation analysis, RMSE values associated with the indices range between 0.35 and 5.93. Also, the Pearson correlation coefficients associated with the indices range between 0.77 and 0.97. Indices with lower RMSE and greater correlation coefficients can provide more accurate geometry degradation predictions for the Melbourne tram network dataset. As represented in this table, OTDI and J Index have lower RMSE compared to the other indices. On the other hands, OTDI and OTGI indices have greater correlation coefficients than the rest of indices. Regarding the performance indicators, the proposed index can be applied in tram track degradation prediction models with strong correlation rate.

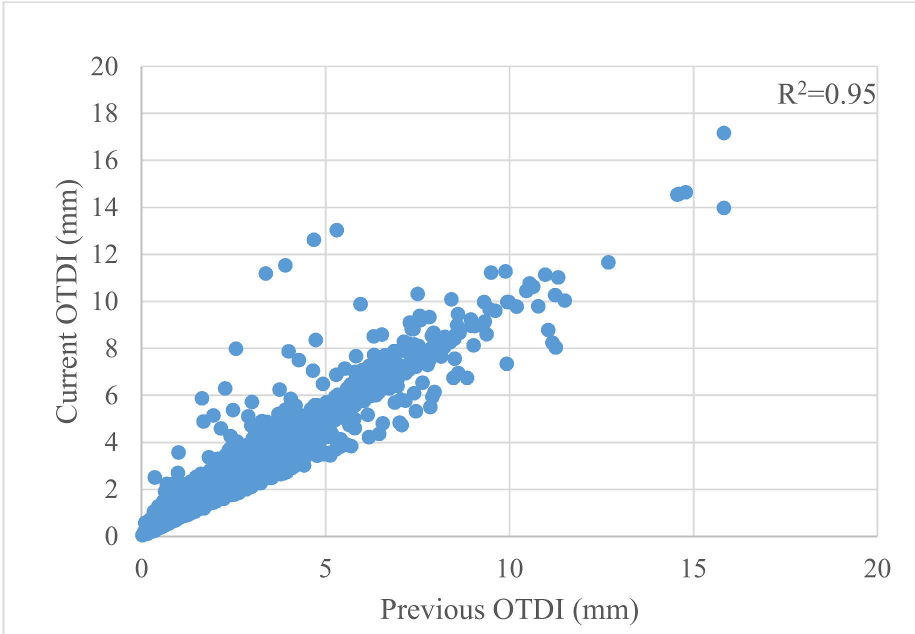

Figure 5.1: The correlation between Previous OTDI and Current OTDI.

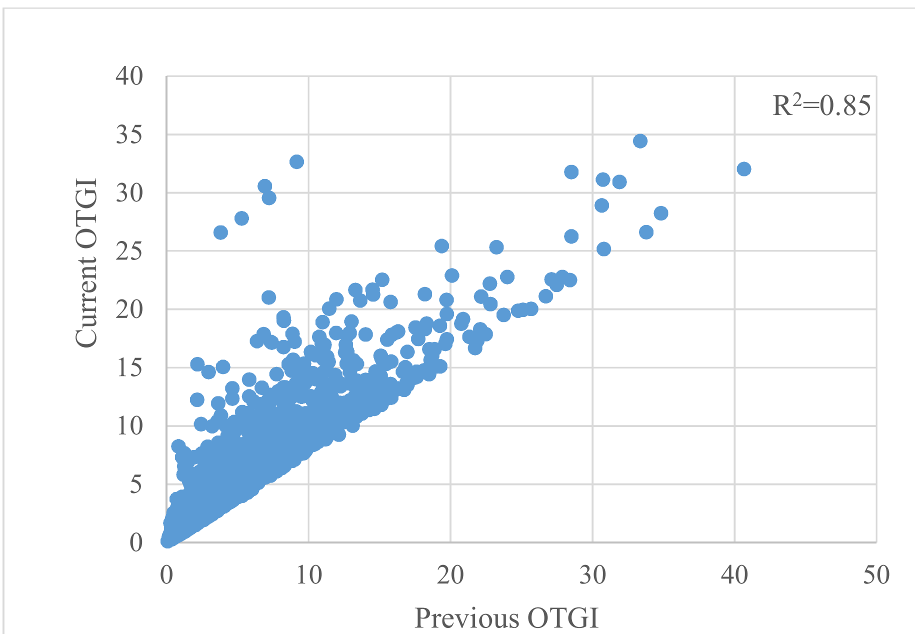

Figure 5.2: The correlation between Previous OTGI and Current OTGI.

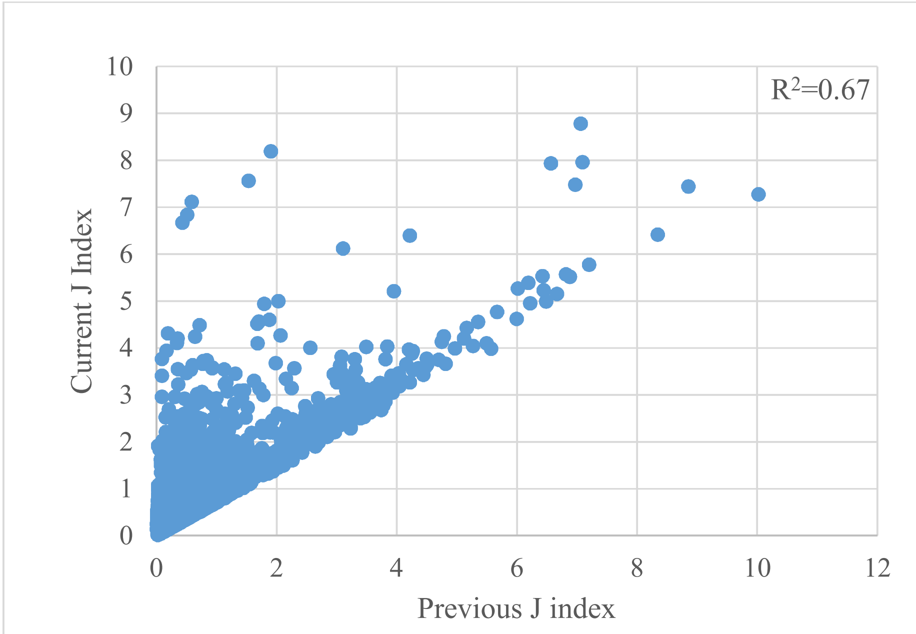

Figure 5.3: The correlation between Previous J-Index and Current J-Index.

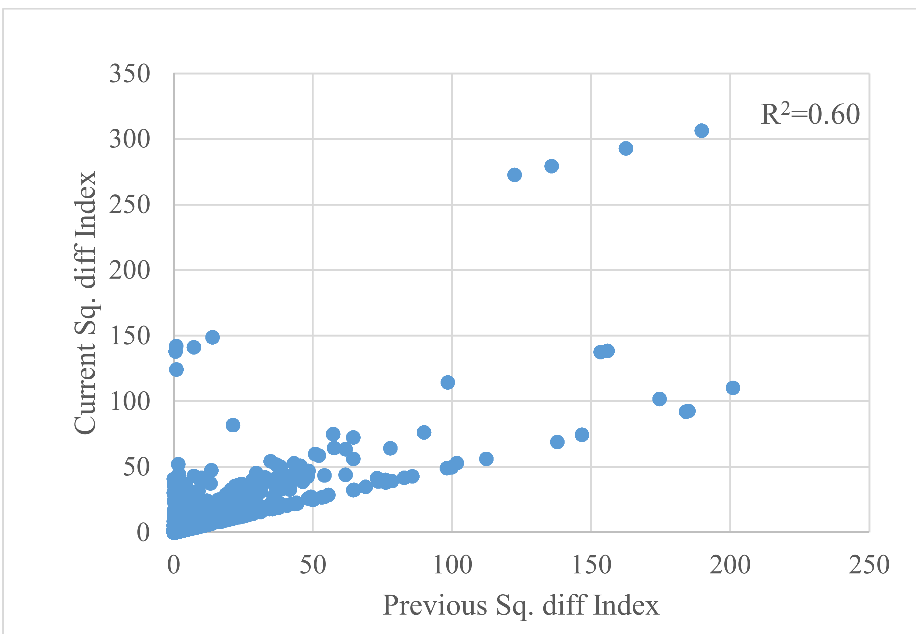

Figure 5.4: The correlation between Previous Amtrak and Current Amtrak.

Table 5.5: The results of the analysis in terms of the Pearson correlation coefficient and RMSE.

Degradation Index	Pearson corr. coef.	Adjusted R^2	RMSE
OTDI	0.97	0.95	0.35
OTGI	0.92	0.85	1.20
J Index	0.82	0.67	0.37
Sq. diff. Index	0.77	0.60	5.93

5.5. Summary

A tram track degradation index is proposed in this chapter based on track geometry parameters included in the research dataset. Track degradation index is a useful measure for infrastructure maintenance management systems as well as prioritising and ranking rail track segments with maintenance needs. To formulate the tram track degradation index and based on the finding of successful degradation indices, the combination of the mean value of geometry deviations and the average of absolute differential geometry deviations were used. In the proposed index, the segments with larger OTDI values are more exposed to degradation and failure than their counterparts with lower OTDI values. The overall track degradation index can examine the condition of track segments comprehensively.

To evaluate and assess the accuracy of the proposed index, the predictability performance of the proposed index along with three major indices were examined. An index which demonstrates greater predictability performance can provide more accurate predictions when is utilised by predictive models. According to the results, the proposed index presents a reasonable correlation with acceptable accuracy compared to other indices. The findings of this chapter indicate that the proposed index can be used as an effective measure for the assessment of the geometric condition of tram tracks as it is easy to develop and apply in the establishment of predictive models based on the previous values of the index. In the following chapter (Chapter 6), the relationship between OTDI and the acceleration data are investigated.

CHAPTER 6

DEVELOPMENT OF DEGRADATION MODELS BASED ON THE ACCELERATION DATA AND TRACK DEGRADATION INDEX

6.1. Introduction

In this chapter, the application of vibration data in predicting tram track geometry deviation is discussed. Although vibration is considered as one of the important factors in passenger ride comfort, yet it has not been applied for predicting tram track degradation in tram network. Currently, vibration data have been used to monitor tram track degradation or in limited context collecting track irregularities in heavy rail network. Vibration can be measured by acceleration signals. The acceleration signal is derived from the movement of railway vehicles on rail structure. In this study, vehicle acceleration data along with other track structural parameters such as track surface and rail support have been used to predict tram track degradation index which has been explained in the previous chapter. For this purpose, Machine Learning (ML) approaches are used for creating the prediction models. Based on the evaluation of the proposed models, the model which provide more accurate predictions on track degradation compared to other developed models is determined. The results of this study can help tram track maintenance management systems to deploy cost-effective maintenance strategies by applying vehicle acceleration data in their decision-making processes. The second section of this chapter presents the current acceleration data application along with their instruments in monitoring and assessment of rail track degradation. The third section presents the procedure for the preparation of a dataset based on track geometry parameters and acceleration data. The fourth section presents the development of models based on acceleration data and track geometry parameters followed by the results and discussion. Section 6, provides the comparison of the two methods of tram track degradation prediction. And the last section presents the summary of this chapter.

6.2. Application of Acceleration Data in Rail Track Monitoring

Vehicle acceleration data have been applied in different rail infrastructure and ride quality studies in the past years. Acceleration data can be measured by different types of inspection devices including Track Vehicle Cars (TRV), trolleys, smartphones, and Condition Monitoring Systems (CMSs).

Acceleration data can then be utilised to assess passenger comfort and rail track quality based on current standards and predefined Track Quality Indices (TQIs). TRVs are inspection rail vehicles designed to capture and store various rail defects as well as track geometry irregularities. These types of vehicles are equipped with different advanced sensors including Non-Destructive Sensors (NDT), accelerometers and optical lasers. Although TRVs are very useful for inspecting rail tracks, there are some drawbacks in relation to their application. First, the cost of running a TRV for inspecting a railway line compared to other devices is high. Secondly, a railway line can be inspected fewer times in a year as the line needs to be closed for train traffic. More information about the application of TRVs in acceleration measurement can be found in Lei (2016), Odashima et al. (2017) and Karis et al. (2018).

Trolleys are cost-effective devices used to manually inspect rail track infrastructure. Due to the lower speed, trolleys cannot be compared with TRVs. Trolleys are more suitable for small-scale rail network inspection. For more details refer to Andani et al. (2018), Gabara and Sawicki (2018) and Evans et al. (2018). Recently, the application of CMSs and smartphones equipped with different sensors has become popular. These devices are cost-effective and accessible. Furthermore, they can be mounted on in-service vehicles without disruptions to train services. Some successful experiences of these devices in rail track monitoring are provided as follows:

Mori et al. (2013) have studied on the development of a compact size on-board device for condition monitoring of railway track. The system comprised of a CPU, microphone for detecting track corrugation, accelerometer (vibration sensors) for detecting track irregularity and a GPS receiver for detecting the position and a computer system for data analysis. This method can estimate riding comfort by determining car-body vibration and evaluates track condition effectively by gathering response characteristics of the car body. Track irregularities can be roughly determined by capturing the value of car body acceleration over the time.

Simonyi et al. (2014) conducted a research to assess various aspects of public transport modes in the city of Budapest with the application of a smartphone and acceleration data. In this research smartphone along with built-in GPS and inertial sensors were used to collect trajectory, velocity and acceleration data. In this research, ISO 2631-5 (2004) has been chosen to assess the effect of vibration on comfort, health, perception and motion sickness. In this regard, the standard defines six different comfort levels including: not uncomfortable (0 m/s^2 - 0.315 m/s^2), a little uncomfortable (0.315 m/s^2 - 0.5 m/s^2), fairly uncomfortable (0.5 m/s^2 - 0.8 m/s^2), uncomfortable (0.8 m/s^2 - 1.25 m/s^2), very uncomfortable (1.25 m/s^2 - 2.0 m/s^2) and extremely uncomfortable (above 2 m/s^2). The experiment of this research has been conducted for public transport systems for three modes of public transportation (a bus ride, a rail track ride and a bus ride on an articulated bus). Vehicle trajectory, speed and longitudinal acceleration data of the vehicles were collected. After conducting the experiment, for assessing the riding comfort of the public transport, the collected acceleration data were compared with the mentioned comfort levels.

Amador-jimenez and Christopher (2016) compared the riding comfort of different modes of public transportation in different cities. In this research vibration as an effective indicator has been chosen for determining the level of comfort and convenience in different modes of public transportation such as buses and trains. For this purpose, smartphones which were mounted to the windscreen were employed for data collection. Their measured values are collected through riding on each specific mode of transportation. Then the data were sorted and filtered to remove unexplained acceleration peak data. Similar to the previous studies, vibration along three axes were examined and weighted using ISO's recommendations. The interaction between rails and car and also the braking systems are among the important factor that can affect the level of acceleration.

Wei et al. (2016) developed a rail track condition monitoring system based on in-service vehicle acceleration measurements. A novel approach for inspecting the track irregularities was presented in this work. For this purpose, different sensors have been mounted on the bogie of a train. These sensors can measures both the lateral acceleration and vertical acceleration of the bogie frame. The track irregularities derived from the system are used to be compared with the predefined thresholds in order to perform track maintenance planning tasks. For the evaluation of the proposed system, the outcomes (track irregularities) were compared with those calculated by a TRV. In general, the proposed track inspection system obtains a very competitive result compared to the TRV.

6.3. Geometry-Acceleration Dataset Preparation

In this chapter, a new dataset based on the combination of track geometry parameters and acceleration data is provided. For this purpose, vehicle acceleration data are collected and provided separately from the track geometry dataset but in the same format involving track record and measurement location. In this research, a CMS has been utilised to collect acceleration data. In order to collect tram acceleration data, accelerometer transducers were installed on the front bogie of in-service trams and aimed to measure lateral acceleration signals. Acceleration data were collected on all the tracks where track geometry parameters were gathered.

By considering the length of 10 m for track segmentation and based on the above data processing, the identification code for the records of the acceleration data have been created. By matching the identification code of the multi-year geometry dataset and the acceleration data, the acceleration-geometry dataset has been developed (Figure 6.1). Before applying the dataset for further processes, error checking needs to be applied to the final dataset to remove errors and incomplete records (in case of one or more elements of a record missed) from the dataset. This process has been undertaken to enhance the dataset and improve the reliability and accuracy of predictions made by models. In Figure 6.2, a sample of the geometry-acceleration dataset is illustrated.

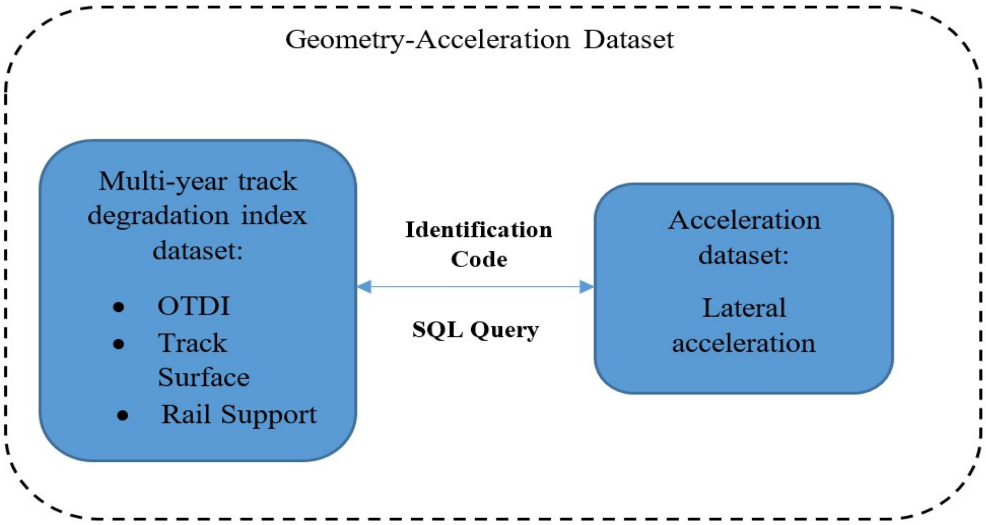

Figure 6.1: Geometry-Acceleration dataset

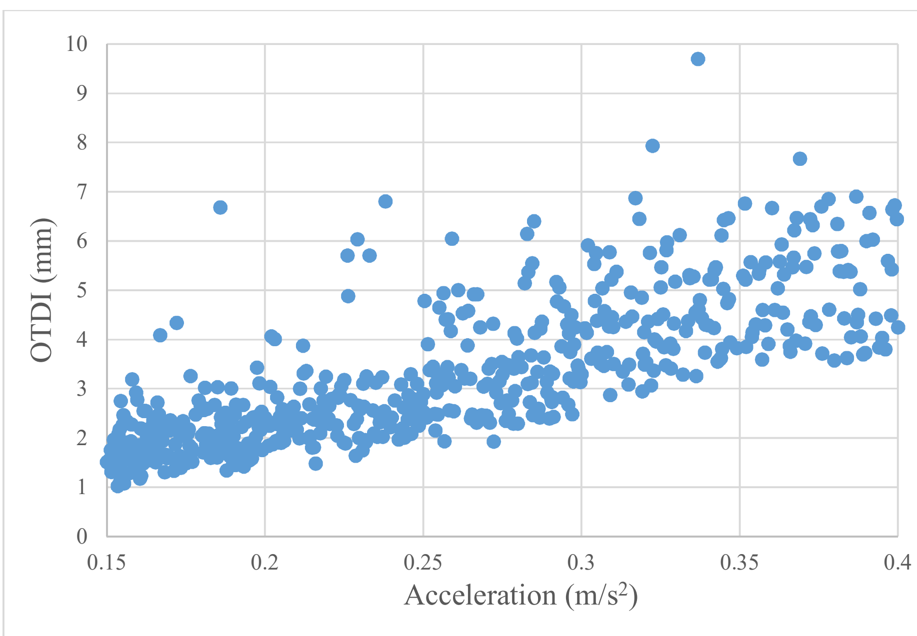

Figure 6.2: A sample of the geometry-acceleration dataset

6.4. Model Development

In this research, three different machine learning models including Random Forest Regression (RFR), SVM and ANN have been developed to predict the degradation index. Explanations about the models have been provided in Section 4 of this research. For the model development, 75% of the dataset were dedicated to training the model and the rest is assigned for data validation and assessing the outcomes. In all the models below, the 5-fold cross-validation is applied to the training dataset to improve the accuracy of the model and avoid over-fitting.

First, ANN models are developed. In this section different ANN models with different architecture in terms of the type of transfer function and the number of neurons in hidden layers are examined. Tangent-Hyperbolicus and Logistic as transfer functions are investigated.

The second set of models are developed based on the RFR approach. In this regard, different models with different *ntree* and *mtry* values are developed. The value of *ntree* ranges between 50 and 100 and the values of *mtry* which is related to the number of the involving parameters (explanatory variables) ranges between 1 and 3.

Lastly, SVM models are developed. In this regard, a plain SVM model and GA-SVM model are examined. In GA-SVM model, the hyper-parameters of the SVM model are optimised with the GA technique to produce better results. For model development, RBF is used as the kernel function.

6.5. Results and Discussion

In this section, the results of the application of the dataset on the proposed models are discussed individually and finally, the comparison of the performances are presented. For evaluation, adjusted R^2, which demonstrates the coefficient of determination the RMSE and MAPE have been used to assess the outcomes of the proposed models as follows:

6.5.1. ANN Model Results

In this section, different models with different variable combinations, number of neurons and transfer function have been examined. The results of the most accurate model in terms of adjusted R^2, the RMSE and MAPE have been tabulated in Table 6.1 below. In Figure 6.2 the correlation between the OTDI data obtained from the analysed geometry deviations and the OTDI predictions predicted by the proposed ANN models is illustrated.

Table 6.1: The results of the ANN model.

Explanatory variables	**Dependent variable**	**Neurons in hidden layers**	**Transfer function**	**Adjusted R^2**	**RMSE**	**MAPE**
Acceleration, Rail support & Track surface	OTDI	4,3	Tanh	0.67	1.13	21.16
		5,3	Tanh	0.66	1.15	21.06
		6,3	Tanh	0.67	1.13	20.75
		7,3	Tanh	0.63	1.20	21.38
		4,3	Logistic	0.67	1.13	20.94
		5,3	Logistic	0.66	1.15	21.07
		6,3	Logistic	0.66	1.14	21.03
		7,3	Logistic	0.65	1.16	21.30

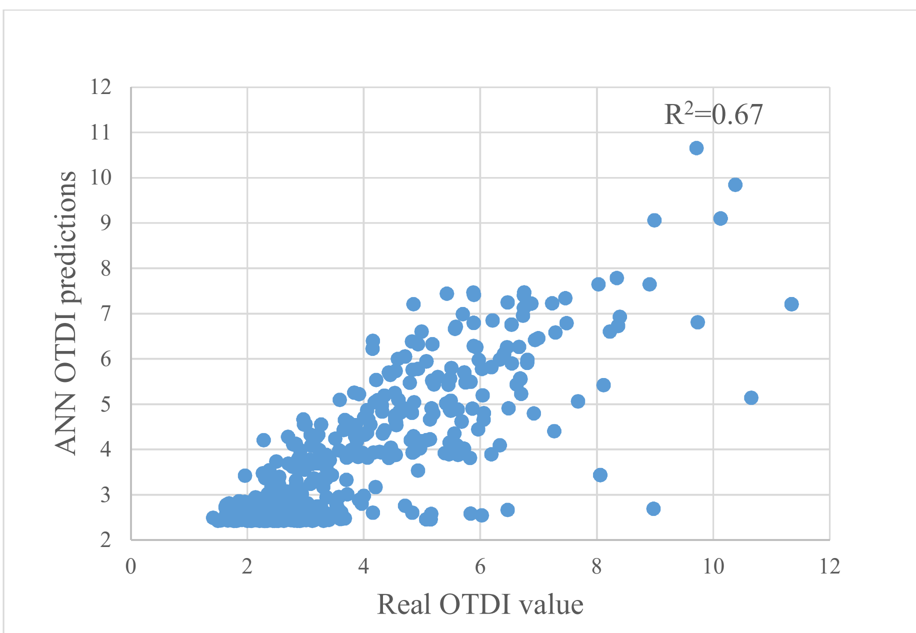

Figure 6.3: Real OTDI values against ANN OTDI predictions.

As illustrated in the above table, different ANN models with regard to the number of neurons in hidden layers and the type of transfer functions have been developed. Adjusted R^2 of the proposed models ranges between 0.63 and 0.67. The RMSE values range between 1.13 and 1.16. The MAPE values range between 20.75 and 21.38. Although the results of the proposed models are very close, an ANN model with number of 6,3 neurons in its hidden layers and Tanh transfer function has provided better prediction compared to other models. Based on these results, an increase in the rate of vehicle acceleration and implementing steel sleeper and asphalt track surface can escalate the rate of degradation in tram tracks. Conversely, the degradation rate is mitigated by implementing concrete sleeper and concrete track surface. The weights of links for the proposed ANN model have been illustrated in Figure 6.3.

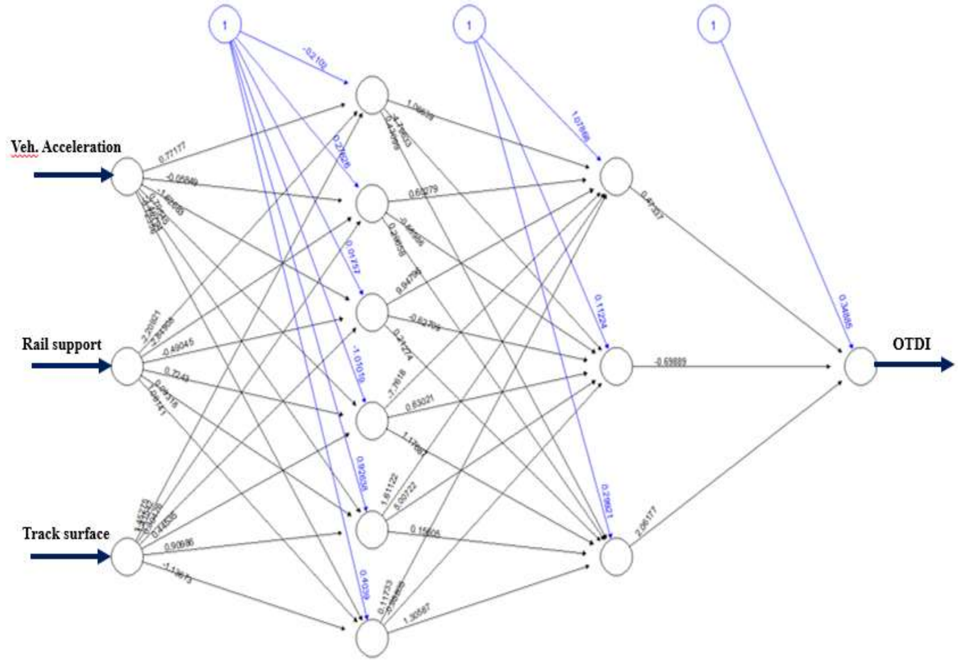

Figure 6.4: Weights of links in the proposed ANN model.

6.5.2. RFR Model Results

In this section, the results of the RFR model development along with the evaluation of the model are provided. and the one with the highest adjusted R^2, the lowest RMSE and MAPE has been selected. The results of the model evaluation are demonstrated in Table 6.2 below. In Figure 6.4 the relationship between the OTDI data obtained from the analysed geometry deviations and the OTDI predictions predicted by the proposed RFR model is depicted.

Table 6.2: The results of the RFR model evaluation.

Explanatory variables	**Dependent variable**	***ntree***	***mtry***	**Adjusted R^2**	**RMSE**	**MAPE**
Acceleration, Rail support & Track surface	OTDI	50	2	0.72	1.05	18.84
		100	2	0.73	1.05	18.81
		50	3	0.73	1.04	16.74
		100	3	0.74	1.04	16.62

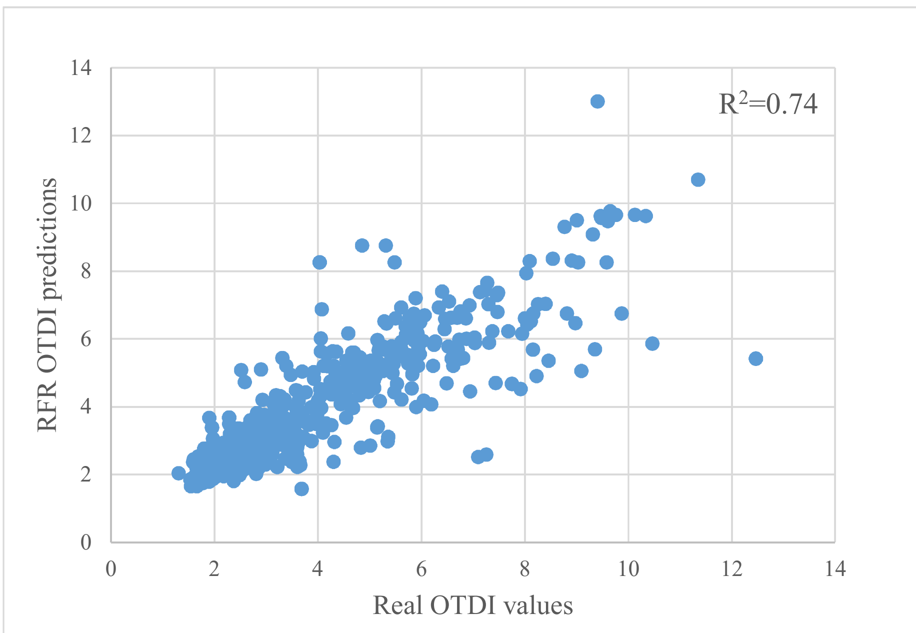

Figure 6.5: Real OTDI values against RFR OTDI predictions.

Based on the results derived from the different models, the model with *mtry* of 3 and *ntree* of 100 has produced slightly more accurate results in terms of adjusted R^2 and the prediction error compared to the other models. As presented in Table 6.2, the proposed RFR model has adjusted R^2 of 0.74, RMSE of 1.04 and MAPE of 16.62. In Figure 6.5, changes in the number of trees versus the OOB error rate has been shown. For the proposed models, the OOB error rate has been decreased gradually by increasing the number of trees. Then the OOB error rate stabilises around the number of 100 trees. Based on the results of the analysis, increasing in the value of vehicle acceleration data leads to greater OTDI which means the higher rate of degradation in tram track. On the other hands, the degradation rate of tram track will be reduced when concrete sleeper and concrete track surface have been used in the tram track infrastructure.

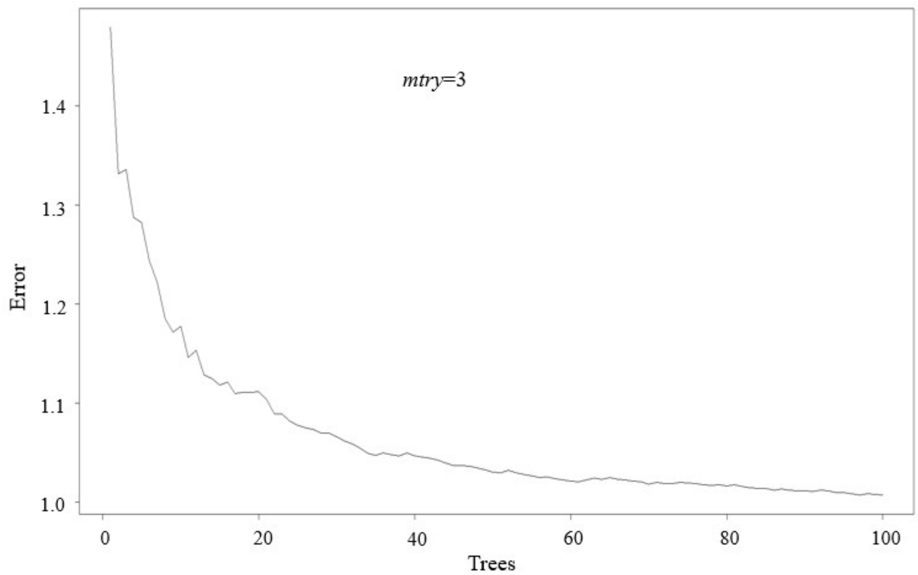

Figure 6.6: OOB error rate of the proposed RFR model.

6.5.3. SVM Model Results

In this section, the results for the SVM model with the highest performance has been demonstrated in Table 6.3 below. As mentioned above, In Figure 6.6 the correlation between the OTDI data obtained from the analysed geometry deviations and the OTDI predictions predicted by the proposed GA-SVM model is illustrated.

Table 6.3: The results of the GA-SVM model evaluation.

Explanatory variables	**Dependent variable**	**Model type**	**Hyper-parameters**	**Kernel function**	**Adjusted R^2**	**RMSE**	**MAPE**
Acceleration, Rail support & Track surface	OTDI	GA-SVM	C= 8.52 γ= 0.41	RBF	0.68	1.13	19.80
		SVM	C=1.00 γ =0.33		0.68	1.14	19.91

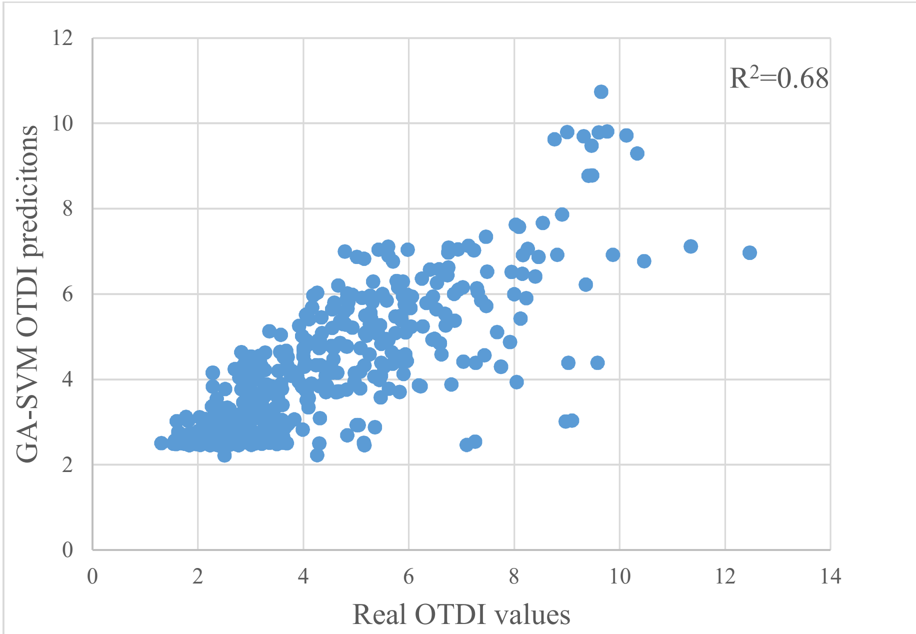

Figure 6.7: Real OTDI values against GA-SVM OTDI predictions.

As presented in Table 6.3 and based on evaluating proposed SVM models, the GA-SVM model has provided slightly better performance to predict the dependent variable compared to the SVM model. For example, for creating the GA-SVM model, the values of tuned hyper-parameters of the proposed model after optimising with GA including C and γ were calculated 8.52 and 0.41, respectively. Based on validating the data, the values of adjusted R^2, RMSE and MAPE for the proposed GA-SVM models are calculated as 0.68, 1.13 and 19.80, respectively. Based on the result of the models, degradation of tram track is worsened by greater vehicle acceleration measured. On the contrary, the degradation rate is decreased by implementing concrete sleeper and concrete track surface.

6.5.4. Comparison of the Proposed Models

In this chapter, three different ML models have been developed and compared. These models were applied to predict the overall track degradation index. The proposed index performs as an indicator of the overall condition of tram tracks based on a combination of two effective parameters in tram track degradation including tram track gauge and twist parameter. According to the results of the evaluation discussed above, the proposed RFR model provided more accurate outcomes in terms of adjusted R^2, the RMSE and MAPE compared to ANN and GA-SVM models which mean that the RFR model predicts the observations more precisely.

Compared to ANN and SVM models, RFR model is computationally less complex but easy to develop. The principles behind this model are more understandable for its users. The model can be optimised by changing the number of grown trees and *mtry*. As shown in the correlation figures, the RFR model is fitted to the dataset more uniformly compared to the other models where the prediction points have mostly gathered at the bottom of the plots.

6.6. Comparison of the Prediction Methods

Similar to heavy rail, conventional tram track maintenance management systems are mainly based on the measurement of track geometry parameters to predict the degradation in tram tracks. As mentioned in previous chapters, measurement of track geometry parameters requires costly works and in some cases imposes delay to the normal function of tram services. In this chapter, a new approach to predict the value of track degradation index (OTDI) were explained which is cost-effective and can be carried out repeatedly without imposing delay to tram services.

In order to compare the method for tram track degradation prediction based on the geometry parameters and the method based on the acceleration data, the performance indicators including the prediction error (RMSE) and coefficient of determination (R^2) have been considered. According to the results of OTDI prediction based on the previous track geometry parameters which have been discussed in the previous chapter, RMSE was 0.35 and R^2 value was 0.95. According to the result of OTDI prediction based on the acceleration data, RMSE was 1.04 and R^2 value was 0.74. Although the prediction error has been increased and R^2 value has been decreased in the latest method, the values of the performance indicators are still in acceptable ranges. This comparison shows that the prediction of tram track degradation based on the acceleration data as a reliable method along with conventional tram track degradation prediction method can be used for maintenance of tram track systems.

6.7. Summary

In this chapter, different approaches of supervised ML models are applied to predict OTDI based on the acceleration data and track structural parameters in a tram network. OTDI is an important parameter to represent the overall condition of tram track segments based on the two geometry parameters including gauge and twist. OTDI as an indicator of tram track degradation can be applied by tram track infrastructure maintenance management systems to prioritise tram track sections for the maintenance operations and renewal purposes.

Acceleration signals can be captured from the movement of rail vehicles on rail tracks. Currently, acceleration data captured form rail track sections are used for the passenger ride comfort classification and health monitoring of rail tracks. In addition to conventional methods and devices to measure rail track irregularities, acceleration data is a useful tool for capturing deviations in rail track parameters. Sudden changes in vehicles acceleration rate can be associated with the presence of rail defects and track irregularities. Integration of acceleration data in the measurement of rail track irregularities can lead to saving the budget and time related to data collection because of using cost-effective CMSs or smartphones. Using acceleration data reduces the maintenance management costs in train and tram infrastructure. Furthermore, by applying this approach on in-service vehicles, condition monitoring of rail track sections can be carried out in several times without imposing delay to the vehicle services.

The models proposed in this chapter can assist rail maintenance engineers in predicting tram OTDI based on the acceleration data without the need for physical measurement of track geometry parameters which can be costly. In this study, the dataset was applied to the RFR, GA-SVM and ANN models. Future OTDI has been considered as the dependent variable and the acceleration data along with track structural parameters have been involved to predict the dependent variables. According to the results of this study, the performance of the proposed models in prediction of OTDI lies within acceptable ranges where the values associated with the coefficient of determination of the models are relatively high and range from 0.64 to 0.74. The errors related to the prediction are close to each other and varies between 1.04 and 1.16. Based on the results, the RFR model can predict future degradation with approximately 10 percent higher R^2 and 9 percent lower prediction error compared to other developed models. In the last section of this chapter, the comparison of the two methods for predicting tram track degradation have been discussed. According to the results, predicting tram track degradation based on the acceleration data as a complement and supportive of conventional methods based on the track geometry parameter can be applied with acceptable accuracy for utilisation in tram track maintenance management systems.

CHAPTER 7

CONCLUSIONS, CONTRIBUTIONS AND FUTURE RESEARCH DIRECTIONS

7.1. Findings

In this research, the impact of different factors in tram track degradation was investigated. After merging track geometry data related to 6 consecutive years of measurements, the multi-year dataset was prepared. For this purpose, different statistical analysis including ANOVA test and Pearson Correlation test was applied to the dataset of the study and this research found:

- Between the existing tram track geometry parameters, gauge and twist deviations have significant contribution to tram track degradation compared to other parameters including Alignment, profile and cross-level. On this basis, previous gauge deviation and previous twist deviation have great contribution accordingly for predicting future gauge and twist deviations.
- Between the existing track structural parameters, track surface and rail support are statistically significant compared to other structural parameters including rail type and rail profile. These variables have reasonable contributions to the degradation of gauge and twist deviations.

In this research different Machine Learning models were developed to predict future gauge and twist deviations based on the previous gauge and twist deviations as well as the structural parameters. This research found that:

- All developed models provided acceptable and reasonable predictions for future gauge and twist deviations which represent their effectiveness in rail track prediction modelling.

- Between developed ML models including SVM, ANN and RFR models, RFR has provided better results in terms of the prediction error and coefficient of determination.
- According to the results, increase in the deviation values of current gauge and twist parameters can result in increase in future gauge and twist deviations.
- According to the result of developed models, concrete track surface and concrete rail support can decrease the rate of degradation and on the contrary, timber rail support and asphalt track surface can increase the rate of track degradation.
- In this research, ANN models which have more neurons in their hidden layers have provided better prediction regarding tram track degradation compared to the ANN models with fewer neurons in their hidden layers.
- In this research, SVM models which their hyper-parameters have been optimised by GA have provided slightly more accurate results compared to SVM models without optimisation.
- RFR model has employed two effective ensemble learning algorithms including Bagging and RSM methods which have boosted the predictions, reduce the over-fitting problem, and improve the performance of the model. In this research RFR models with higher values of *ntree* and *mtry* have provided more accurate results compared to other RFR models.
- The comparison of the developed models showed that regarding gauge deviating prediction, RFR model provided R^2, 6 percent higher than GA-SVM model and about 11 percent higher than the ANN model. Also RFR model provided RMSE, more than 16 percent lower than GA-SVM model and more than 42 percent lower than the ANN model.
- Regarding twist deviation prediction, RFR model provided R^2, more than 20 percent higher than GA-SVM and ANN model. Also RFR model provided RMSE, 22 percent lower than GA-SVM model and about 16 percent lower than the ANN model.

In this research index development for tram track degradation has been introduced. Different track geometry degradation indices have been reviewed and the different formulation has been examined. Effective elements and parameters in developing a degradation index which can increasingly improve the performance of the proposed index have been investigated. Current indices are mostly developed and designed based on the heavy rail track geometry parameters. The summary of findings of the index development have been presented as follows:

- The sum of the absolute value of differential geometry deviations of two consecutive track geometry parameters (gauge and twist) and the mean value of track geometry parameters have been included in the formulation for tram track degradation index.
- Application of the proposed index on the research dataset showed that the index can effectively reflect the changes occurred in gauge and twits deviations over time which can affect the functionality of tram track segments
- As track degradation indices are used to predict the future condition of rail tracks, the predictability performance of indices as an important factor has been utilised in the evaluation of the existing indices and the proposed index. For examining the predictability performance, the correlation between previous and future geometry deviations (gauge and twist) have been investigated for four common indices including the proposed index. It has been revealed that the proposed index provides a more accurate correlation compared to the others.
- According to the results of the predictability performance, OTDI provided R^2, 28 percent higher than *J* index, 10 percent higher than OTGI and 37 percent higher than Amtrak index.

In this research, the relationship between ride comfort data and tram track degradation index has been investigated and different models to predict the index based on the acceleration data have been developed. The finding can be summarised as below.

- All developed models provided acceptable and reasonable predictions for the tram track degradation index based on the acceleration data which represent their accuracy in the modelling of geometry/acceleration dataset.

- According to the results, increase in the amount of vehicle acceleration data which can be expressed by lateral vibration can result in increase in the value of track degradation index. In other words, significant lateral vibration indicates the deviations in the tram track.
- According to the result, concrete track surface and concrete rail support can decrease the rate of the predicted degradation index and on the contrary, timber rail support and asphalt track surface can increase the rate of the predicted index.
- Among three different machine learning models, RFR model provided better prediction compared to other models.
- According to the results, the RFR model can predict future degradation with more than 9 percent higher R^2 than ANN and GA-SVM models. Also, the RFR model can predict future degradation with about 8 percent lower RMSE than ANN and GA-SVM models. Similarly, the RFR model can predict with more than 16 percent lower MAPE than the other models.

7.2. Contributions

In terms of track geometry degradation prediction modelling based on the geometry dataset and track structural parameters, the contribution of this research as follows.

- Most studies conducted in the field of track degradation prediction modelling have dealt with heavy rail degradation process while tram track degradation modelling was neglected and a limited number of researches have been conducted in this area.
- Regression models were common models used for modelling degradation process in tram track in previous studies while the degradation process in tram track demands more elaborate degradation prediction models. In this research, different types of machine learning models were used for improving the effectiveness of the degradation prediction models.

In terms of the development of tram track degradation index, the contribution of this research can be outlined as follows.

- Existing track degradation indices are developed based on the heavy rail track geometry data. In this research tram track degradation index based on the tram track dataset has been proposed.
- In this research, by examining literature review a new formulation for developing a track degradation index has been provided which can be applied to both heavy rail and tram track datasets. All existing degradation indices are designed and developed for heavy rail tracks.
- Predictability performance as an important factor for evaluating different degradation indices has been introduced in this paper. By examining the predictability performance, the index which has provided better predictions in future has been determined.

In terms of the development of a degradation prediction model based on the acceleration data and the track degradation index the contribution of this research can be summarised as follows:

- Acceleration data have been used in several studies to reflect and monitor track geometry deviation in rail track inspection and maintenance practices. In this study, acceleration data have been integrated into tram track degradation prediction modelling to predict the overall track degradation index.
- The proposed model based on the acceleration data can help rail track maintenance engineers to calculate overall degradation index without physical measurement of track geometry parameters. This process can lower the maintenance cost and expenditure for conducting periodical rail track inspection and monitoring. Consequently, tram tracks can be inspected in shorter intervals.

7.3. Future Research Directions and Recommendations

- It would be worthwhile to apply noise-cancelling methods such as Locally Estimated Scatterplot Smoothing (LOESS), Spline Smoothing, Exponential

Smoothing and Moving Average to acceleration datasets to increase the accuracy of the dataset and remove noisy data from the model development.

- The result of the model development showed that ensemble learning methods have provided better predictions on tram track data compared to the other models. By applying ensemble techniques, several learning datasets can be produced randomly which can increase the performance of the prediction models. It would be worthwhile to develop more models based on the ensemble learning methods such as Boosting and Bayesian model averaging.
- As the contributing parameters in tram track degradation can differ in different datasets, the application of this index to the different dataset can provide new results. Increase in the number of contributing parameters in the index development process can directly improve the comprehensiveness of the proposed index.
- The Methodology developed in this research can be applied to other datasets by examining and investigating the impact of different types of acceleration data including vertical and longitudinal acceleration on track degradation index.

References

Ahac, M., & Lakusic, S. (2017). Track Gauge Degradation Modelling on Small Urban Rail Networks: Zagreb Tram System Case Study. *Urban Transport Systems*, 1.

Alemazkoor, N., Ruppert, C. J., & Meidani, H. (2015). Track Geometry Analytics. *RAS Problem Solving Competition Report*. Mumbai, India.

Altman, N., & Krzywinski, M. (2017). Points of significance: Ensemble methods: bagging and random forests. *Nature Methods*, 14(10), 933–934.

Amador-jimenez, L., & Christopher, A. (2016). A Comfort index for public transportation : Case study of Montreal. *IEEE International Conference on Intelligent Transportation Engineering* (pp. 3–7).

Andani, M., Mohammed, A., Jain, A., & Ahmadian, M. (2018). Application of LIDAR technology for rail surface monitoring and quality indexing. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(5), 1398–1406.

Andrade, A. R., & Teixeira, P. F. (2012). A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, 36(1), 1–8.

Andrade, A. R., & Teixeira, P. F. (2012). A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, 36(1), 1–8.

Andrade, A. R., & Teixeira, P. F. (2014). Unplanned-maintenance needs related to rail track geometry. *Proceedings of the Institution of Civil Engineers - Transport*, 167(6), 400–410.

Andrade, A. R., & Teixeira, P. F. (2015). Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models. *Reliability Engineering & System Safety*, 142, 169–183.

Andrews, J., Prescott, D., & De Rozières, F. (2014). A stochastic model for railway track asset management. *Reliability Engineering and System Safety*, 130, 76–84.

Asada, T., Roberts, C., & Koseki, T. (2013). An algorithm for improved performance of railway condition monitoring equipment: Alternating-current point machine case study. *Transportation Research Part C: Emerging Technologies*, 30, 81–92.

Audley, M., & Andrews, J. D. (2013). The effects of tamping on railway track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4), 376–391.

Bai, L., Liu, R., Sun, Q., Wang, F., & Xu, P. (2015). Markov-based model for the prediction of railway track irregularities. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 229(2), 150–159.

Baldi, M. M., Heinicke, F., Simroth, A., & Tadei, R. (2016). New heuristics for the Stochastic Tactical Railway Maintenance Problem. *Omega*, 63, 94–102.

Berawi, A. R. B. (2013). *Improving railway track maintenance using power spectral density (PSD)*. Universidade do Porto.

Caetano, L. F., & Teixeira, P. F. (2015). Optimisation model to schedule railway track renewal operations: a life-cycle cost approach. *Structure and Infrastructure Engineering*, 11(11), 1524–1536.

De Bruijn, H., & Veeneman, W. (2009). Decision-making for light rail. *Transportation Research Part A: Policy and Practice*, *43*(4), 349-359.

DeGroot, M. H., & Schervish, M. J. (2012). *Probability and statistics*. Pearson Education.

Dell'Orco, M., Ottomanelli, M., Caggiani, L., & Sassanelli, D. (2008). New Decision Support System for Optimization of Rail Track Maintenance Planning Based on Adaptive Neurofuzzy Inference System. Transportation Research Record: *Journal of the Transportation Research Board*, 2043, 49–54.

Demharter, K. (1982). Setzungsverhalten des gleisrostes unter vertikaler lasteinwirkung. Deutschland.

Diemer, M. J., Currie, G., De Gruyter, C., & Hopkins, I. (2018). Filling the space between trams and place: Adapting the 'Movement & Place' framework to Melbourne's tram network. *Journal of Transport Geography*, 70, 215–227.

Dincer, I., Hogerwaard, J., & Zamfirescu, C. (2015). *Clean rail transportation options, Green energy and technology*. Oshawa, Canada: Springer International Publishing.

European Standard Standardization Committee (2010). Railway applications-Track geometry quality - Part 5: Geometry quality levels - Plain line.

Evans, M., Lucas, A., & Ingram, I. (2018). The inspection of level crossing rails using guided waves. *Construction and Building Materials*, 179, 614–618.

Fagerland, M. W., & Hosmer, D. W. (2017). How to test for goodness of fit in ordinal logistic regression models. *The Stata Journal*, 17(3), 668–686.

Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2017). Rail degradation predication: Melbourne tram system case study. *Australian Transport Research Forum (ATRF).* Auckland, New Zealand.

Falamarzi, A., Moridpour, S., Nazem, M., & Cheraghi, S. (2018b). Development of a random forests regression model to predict track degradation index: Melbourne case study. *Australian Transport Research Forum (ATRF)*. Darwin, Australia.

Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2018a). Rail degradation prediction models for tram system: Melbourne case study. *Journal of Advanced Transportation*, 2018.

Falamarzi, A., Moridpour, S., Nazem, M., & Hesami, R. (2019). Integration of genetic algorithm and support vector machine to predict rail track degradation. *MATEC Web of Conferences*, 259.

Fang, Z., Cheng, Q., Jia, R., & Liu, Z. (2018). Urban rail transit demand analysis and prediction: A review of recent studies. *Intelligent Interactive Multimedia Systems and Services*, 98, 300-309.

Gabara, G., & Sawicki, P. (2018). A New approach for inspection of selected geometric parameters of a railway track using image-based point clouds. *Sensors*, 18(791), 16.

Gaudry, M., Lapeyre, B., & Quinet, É. (2016). Infrastructure maintenance, regeneration and service quality economics: A rail example. *Transportation Research Part B: Methodological*, 86, 181-210

D'Angelo, G., Bressi, S., Giunta, M., Presti, D. L., & Thom, N. (2018). Novel performance-based technique for predicting maintenance strategy of bitumen stabilised ballast. *Construction and Building Materials*, 161, 1-8.

Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P., & Sun, Y. (2010). A review on degradation models in reliability analysis. *Engineering Asset Lifecycle Management* (pp. 369–384). Athens, Greece.

Guler, H. (2013). Decision Support System for Railway Track Maintenance and Renewal Management. *Journal of Computing in Civil Engineering*, 27(3), 292–306.

Guler, H. (2014). Prediction of railway track geometry degradation using artificial neural networks: a case study for Turkish state railways. *Structure and Infrastructure Engineering*, 10(5), 614-626.

Guler, H., Jovanovic, S., & Evren, G. (2011). Modelling railway track geometry deterioration. *Proceedings of the Institution of Civil Engineers - Transport*, 164(2), 65–75.

Hajibabai, L., Saat, M. R., & Ouyang, Y. (2012). Wayside Defect Detector Data Mining to Predict Potential WILD Train Stops. *Proceedings of American Railway Engineering and Maintenance-of-Way Association Annual Meeting*. Chicago, Illinois.

He, Q., Li, H., Bhattacharjya, D., Parikh, D. P., & Hampapur, A. (2013). Railway Track Geometry Defect Modeling: Deterioration, Derailment Risk and Optimal Repair. *Proceedings of the Transportation Research Board Annual Meeting*. The Academy of Transportation Research Board.

He, Q., Li, H., Bhattacharjya, D., Parikh, D. P., & Hampapur, A. (2015). Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. *Journal of the Operational Research Society*, 66(3), 392-404.

Higgins, C., & Liu, X. (2018). Modeling of track geometry degradation and decisions on safety and maintenance: A literature review and possible future research directions. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(5), 1385-1397.

Hosmer Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression.* Wiley series in probability and statistics.

Hua, J., Shen, Z., & Zhong, S. (2017). We can track you if you take the metro: Tracking metro riders using accelerometers on smartphones. *IEEE Transactions on Information Forensics and Security*, 12(2), 286–297.

Jamshidi, A., Faghih Roohi, S., Núñez, A., Babuska, R., De Schutter, B., Dollevoet, R., & Li, Z. (2016). Probabilistic Defect-Based Risk Assessment Approach for Rail Failures in Railway Infrastructure. *International Federation of Automatic Control*, 49(3), 73–77.

Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., & De Schutter, B. (2017). A big data analysis approach for rail failure risk assessment. *Risk analysis*, 37(8), 1495-1507.

Jeong, M. C., Lee, S. J., Cha, K., Zi, G., & Kong, J. S. (2019). Probabilistic model forecasting for rail wear in Seoul metro based on Bayesian theory. *Engineering Failure Analysis, 96,* 202–210.

Jia, C., Xu, W., Wang, F., & Wang, H. (2012). Track irregularity time series analysis and trend forecasting. *Discrete Dynamics in Nature and Society*, 2012.

Jiang, Z., Gu, J., Han, Y., Fan, W., & Chen, J. (2018). Modeling actual dwell time for rail transit using data analytics and support vector regression. *Journal of Transportation Engineering, Part A: Systems,* 5(11), 1–12.

Jovanović, S., Guler, H., & Čoko, B. 2015. Track degradation analysis in the scope of railway infrastructure maintenance management systems. *Gradjevinar,* 67(3), 247–258.

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data analysis: A model comparison approach to regression, ANOVA, and beyond*. New York, USA: Routledge.

Karimpour, M., Hitihamillage, L., Elkhoury, N., Moridpour, S., & Hesami, R. (2018). Fuzzy Approach in Rail Track Degradation Prediction. *Journal of Advanced Transportation*, 2018.

Karis, T., Berg, M., Stichel, S., Li, M., Thomas, D., & Dirks, B. (2018). Correlation of track irregularities and vehicle responses based on measured data. *Vehicle System Dynamics*, 56(6), 967–981.

Knowles, R. D., & Ferbrache, F. (2016). Evaluation of wider economic impacts of light rail investment on cities. *Journal of Transport Geography*, 54, 430-439.

Kopf, F., Maras, I., Gasser, F., Norkauer, A., & Ritz, O. (2009). *Cost-effective track maintenance, renewal & refurbishment methods*. Belgium.

Kwon, H. B. (2017). Exploring the predictive potential of artificial neural networks in conjunction with DEA in railroad performance modeling. *International Journal of Production Economics*, 183(November 2016), 159–170.

Laryea, S., Baghsorkhi, M. S., Ferellec, J. F., McDowell, G. R., & Chen, C. (2014). Comparison of performance of concrete and steel sleepers using experimental and discrete element methods. *Transportation Geotechnics*, 1(4), 225–240.

Lasisi, A., & Attoh-Okine, N. (2018). Principal components analysis and track quality index: A machine learning approach. *Transportation Research Part C: Emerging Technologies*, 91, 230–248.

Lei, X. (2016). *High speed railway track dynamics models, algorithms and applications: Track dynamics research contents and related standards*. Nanchang: Springer Press Beijing.

Li, F. Y., Xu, Y. D., Li, H. F., Shen, J. F., Liu, W. Y., & Qiu, J. X. (2016). A comparative study of the TQI method and process performance index method in the quality evaluation of track fine adjustment. *Proceedings of the 15th International Conference on Railway Engineering Design and Operation*, 162, 33–40.

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies*, 45, 17–26.

Li, W., Pu, H., Schonfeld, P., Yang, J., Zhang, H., Wang, L., & Xiong, J. (2017). Mountain railway alignment optimization with bidirectional distance transform and genetic algorithm. *Computer-Aided Civil and Infrastructure Engineering*, 32(8), 691–709.

Liu, K., Wang, M., Cao, Y., Zhu, W., & Yang, G. (2018). Susceptibility of existing and planned Chinese railway system subjected to rainfall-induced multi-hazards. *Transportation Research Part A: Policy and Practice*, 117, 214–226.

Liu, R. K., Xu, P., Sun, Z. Z., Zou, C., & Sun, Q. X. (2015). Establishment of track quality index standard recommendations for Beijing metro. *Discrete Dynamics in Nature and Society*, 2015.

Madejski, J., & Grabczyk, J. (2002). Continuous geometry measurement for diagnostics of tracks and switches. *Proceedings of the International Conference on Switches*.

Marković, N., Milinković, S., Tikhonov, K. S., & Schonfeld, P. (2015). Analyzing passenger train arrival delays with support vector regression. *Transportation Research Part C: Emerging Technologies*, 56, 251–262.

Mercier, S., Meier-Hirmer, C., & Roussignol, M. (2012). Bivariate Gamma wear processes for track geometry modelling, with application to intervention scheduling. *Structure and Infrastructure Engineering*, 8(4), 357–366.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. 2013. *Machine Learning: An Artificial Intelligence Approach*. Springer Science & Business Media. Berlin, Heidelberg.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis (Fifth)*. Canada: Wiley.

Morant, A., Larsson-Kraik, P. O., & Kumar, U. (2016). Data-driven model for maintenance decision support: A case study of railway signalling systems. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 230(1), 220–234.

Mori, H., Sato, Y., Ohno, H., Tsunashima, H., & Saito, Y. (2013). Development of compact size onboard device for condition monitoring of railway tracks. *Journal of Mechanical Systems for Transportation and Logistics*, 6(2), 142–149.

Moridpour, S., & Hesami, R. (2015). Degradation and performance specification of Melbourne tram tracks. *ICTIS 2015 - 3rd International Conference on Transportation Information and Safety* (pp. 270–276).

Moridpour, S., Mazloumi, E., & Hesami, R. (2017). Application of Artificial Neural Networks in Predicting the Degradation of Tram Tracks Using Maintenance Data. *Applied Big Data Analytics in Operations Management* (pp. 30–54).

Mundrey, J. S. (2009). *Railway track engineering (Fourth edition)*. New Delhi: Tata McGraw-Hill Education.

Naznin, F., Currie, G., & Logan, D. (2016). Exploring the impacts of factors contributing to tram-involved serious injury crashes on Melbourne tram routes. *Accident Analysis and Prevention*, 94, 238–244.

Nicodeme, C., Stanciulescu, B., Onera, R. C., Nicodeme, C., & Fassi, S. El. (2017). Wheel-rail contact analysis system using spectral signatures for train automation and traffic management. *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)* (pp. 347–351).

Nunez, A., Hendriks, J., Li, Z., De Schutter, B., & Dollevoet, R. (2014). Facilitating maintenance decisions on the Dutch railways using big data: The ABA case study. *IEEE International Conference on Big Data* (pp. 48–53).

Odashima, M., Azami, S., Naganuma, Y., Mori, H., & Tsunashima, H. (2017). Track geometry estimation of a conventional railway from car-body acceleration measurement. *The Japan Society of Mechanical Engineers*, 4(1), 16-00498.

Odolinski, K., & Smith, A. S. J. (2016). Assessing the cost impact of competitive tendering in rail infrastructure maintenance services: Evidence from the swedish reforms (1999 to 2011). *Journal of Transport Economics and Policy*, 50(1), 93–112.

Odolinski, K., & Smith, A. S. J. (2016). Assessing the cost impact of competitive tendering in rail infrastructure maintenance services: Evidence from the Swedish reforms (1999 to 2011). *Journal of Transport Economics and Policy*, 50(1), 93–112.

Paneiro, G., Durão, F. O., Costa e Silva, M., & Neves, P. F. (2018). Artificial neural network model for ground vibration amplitudes prediction due to light railway traffic in urban areas. *Neural Computing and Applications*, 29(11), 1045–1057.

Prescott, D., & Andrews, J. (2013). A track ballast maintenance and inspection model for a rail network. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability,* 227, 251–266.

PTV (2018). *Annual report 2017-2018*, Public Transport Victoria, Melbourne.

Quiroga, L., & Schnieder, E. (2010). Modelling of high speed railroad geometry ageing as a discrete-continuous process. Stochastic *Modeling Techniques and Data Analysis International Conference* (pp. 1–8).

Rana, A., Verma, A. K., & Srividya, A. (2016). Maintenance of Large Engineering Systems. *Current Trends in Reliability, Availability, Maintainability and Safety*, 599–609.

Rashidi, M., Samali, B., & Sharafi, P. (2016). A new model for bridge management: Part B: decision support system for remediation planning. *Australian Journal of Civil Engineering*, 14(1), 46–53.

Roghani, A. (2017). *A quantitative evaluation of the impact of soft subgrades on railway track structure*. University of Alberta.

Sadeghi, J. (2010). Development of Railway Track Geometry Indexes Based on Statistical Distribution of Geometry Data. *Journal of Transportation Engineering*, 136(8), 693–700.

Sadeghi, J. (2010). Development of Railway Track Geometry Indices Based on Statistical Distribution of Geometry Data. *Journal of Transportation Engineering*, 136(8), 693–700.

Sadeghi, J., & Askarinejad, H. (2010). Development of improved railway track degradation models. *Structure and Infrastructure Engineering*, 6(6), 675–688.

Sadeghi, J., & Askarinejad, H. (2012). Application of neural networks in evaluation of railway track quality condition. *Journal of Mechanical Science and Technology*, 26(1), 113–122.

Salvador, P., Naranjo, V., Insa, R., & Teixeira, P. (2016). Axlebox accelerations: Their acquisition and time-frequency characterisation for railway track monitoring purposes. Measurement: *Journal of the International Measurement Confederation*, 82(518), 301–312.

Santos, R., Teixeira, P. F., & Antunes, A. P. (2015). Planning and scheduling efficient heavy rail track maintenance through a Decision Rules Model. *Research in Transportation Economics*, 54, 20–32.

Sarstedt, M., & Mooi, E. (2019). *A Concise Guide to Market Research (Third)*. Germany: Springer.

Sato, Y. (1995). Japanese Studies on Deterioration of Ballasted Track. *Vehicle System Dynamics*, 24, 197–208.

Setiawan, D., & Rosyidi, S. A. (2016). Track quality index as track quality indicator. *The 19th International Symposium of FSTPT* (pp. 1–10).

Shafahi, Y., Masoudi, P., & Hakhamaneshi, R. (2008). Track Degradation Prediction Models, Using Markov Chain, Artificial Neural and Neuro-Fuzzy Network. *8th World Congress on Railway Research* (pp. 1–9). Seoul, Korea.

Sharma, S., Cui, Y., He, Q., Mohammadi, R., & Li, Z. (2018). Data-driven optimization of Railway maintenance for track geometry. *Transportation Research Part C: Emerging Technologies*, 90, 34-58.

Shenton, M. J. (1985). Ballast deformation and track deterioration. *Track Technology*, 253–265.

Simonyi, E., Fazekas, Z., & Gáspár, P. (2014). Smartphone application for assessing various aspects of urban public transport. *Transportation Research Procedia,* 3, 185–194.

Soleimanmeigouni, I., Ahmadi, A., & Kumar, U. (2018). Track geometry degradation and maintenance modelling: A review. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(1), 73–102.

Thaduri, A., Galar, D., & Kumar, U. (2015). Railway assets: A potential domain for big data analytics. *Procedia Computer Science*, 53, 457–467.

Vale, C., & M. Lurdes, S. (2013). Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line. *Reliability Engineering and System Safety*, 116, 91–98.

Van Der Westhuizen, N. J., & Grabe, P. J. (2013). The integration of railway asset management information to ensure maintenance effectiveness. *Journal of the South African Institution of Civil Engineering*, 55(3), 18–29.

Wang, Y., Cheng, H., & Li, S. (2018). Passenger flow forecast model for intercity high speed railway - A neural network-based analysis. *Journal of Interdisciplinary Mathematics*, 21(4), 897–906.

Wei, X., Liu, F., & Jia, L. (2016). Urban rail track condition monitoring based on in-service vehicle acceleration measurements. *Measurement*, 80(2016), 217–228.

Were, K., Bui, D. T., Dick, O. B., & Singh, B. R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators*, 52, 394–403.

Westgeest, F., Dekker, R., & Fischer, R. (2012). Predicting rail geometry deterioration by regression models. *Advances in Safety, Reliability and Risk Management* (pp. 926–933).

Wilson, A., & Kerr, M. (2013). *Rail installation and repair*. Sydney.

Xu, P., Sun, Q., Liu, R., & Wang, F. (2011). A short-range prediction model for track quality index. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 225(3), 277–285.

Yadav, N., Yadav, A., & Kumar, M. (2015). *An Introduction to Neural Network Methods for Differential Equations*. London: Springer.

Yousefikia, M., Moridpour, S., Setunge, S., & Mazloumi, E. (2014). Modeling Degradation of Tracks for Maintenance Planning on a Tram Line. *Journal of Traffic and Logistics Engineering*, 2(2).

Zhang, Y., Kim, C. W., & Tee, K. F. (2017). Maintenance management of offshore structures using Markov process model with random transition probabilities. *Structure and Infrastructure Engineering*, 13(8), 1068–1080.

Zhu, M., Cheng, X., Miao, L., Sun, X., & Wang, S. (2013). Advanced stochastic modeling of railway track irregularities. *Advances in Mechanical Engineering*, 2013.

Zhu, Q. & Taher, A. (2015). Studies in Fuzziness and Soft Computing Complex System *Modelling and Control through Intelligent Soft Computations*. Switzerland: Springer International Publishing.

Zimmermann, H. J. (2010). *Fuzzy set theory*. Wiley Interdisciplinary Reviews: Computational Statistics.